Towards Ultrafast Communications: Nonlinear Coupling Dynamics & Light-Semiconductor Interaction

Wang, W.

2007

citation for published version (APA)
Wang, W. (2007). *Towards Ultrafast Communications: Nonlinear Coupling Dynamics & Light-Semiconductor Interaction*. [PhD-Thesis - Research and graduation internal, Vrije Universiteit Amsterdam].

Towards Ultrafast Communications:
Nonlinear Coupling Dynamics &
Light-Semiconductor Interaction

VRIJE UNIVERSITEIT

Towards Ultrafast Communications: Nonlinear Coupling Dynamics & Light-Semiconductor Interaction

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad van Doctor aan
de Vrije Universiteit Amsterdam,
op gezag van de rector magnificus
prof.dr. L.M. Bouter,
in het openbaar te verdedigen
ten overstaan van de promotiecommissie
van de faculteit der Exacte Wetenschappen
op maandag 19 maart 2007 om 15.45 uur
in het auditorium van de universiteit,
De Boelelaan 1105

door

Wenfeng Wang

geboren te Hubei, China

promotor:	prof.dr. D. Lenstra
copromotor:	dr. K. Allaart

This work was part of the project ”Towards Ultrafast Communications”, which belongs to the Freeband Communication Impulse of the Technology Program of the Netherlands’ Ministry of Economic Affairs.

Contents

1 Introduction 1

1.1 Nonlinear propagation of light 2

1.2 Light-semiconductor interaction 4

1.3 Outline of the thesis 6

2 Manipulation of ultrashort optical pulses 7

2.1 Solving coupled nonlinear Schrödinger Equations 8

2.2 New normalization 11

2.3 Ultrashort rectangular optical pulse generation 14

2.3.1 Introduction 14

2.3.2 Theory 15

2.3.3 Generation of a flat-top pulse 15

2.3.4 Generation of a rectangular pulse 16

2.3.5 Conclusion 23

2.4 All-optical limiting 23

2.4.1 Introduction 23

2.4.2 Theory 25

2.4.3 CW analytical solution 26

2.4.4 Numerical analysis 28

2.4.5 Conclusion 35

3 Optical signal processing by semiconductor optical amplifiers 37

3.1 Electronic structure of a semiconductor 38

3.1.1 The $\vec{k} \cdot \vec{p}$ method 38

3.1.2 Truncation of the space to X, Y, Z, S 40

3.1.3 Inclusion of the interaction with remote states 41

3.1.4 Spin-orbit interaction 43

3.1.5 Wave functions and energy bands 45

3.1.6 Uniaxially strained semiconductor 48

3.2 Optical properties of a semiconductor . . . 50
3.2.1 Interaction of light and matter . . . 50
3.2.2 Bulk semiconductor . . . 53
3.2.3 Uniaxially strained semiconductor . . . 56
3.3 Semiconductor optical amplifier (SOA) . . . 56
3.3.1 Gain and refractive index . . . 56
3.3.2 Anisotropic gain of a SOA . . . 60
3.4 Optical signal processing (OSP) based on SOA . . . 63
3.4.1 Mechanisms of OSP . . . 63
3.4.2 Ultrafast nonlinearities of a SOA . . . 64
3.5 Summary and Discussion . . . 66

4 Optical polarization in electron-spin polarized semiconductors 67
4.1 Introduction . . . 67
4.2 Model . . . 68
4.3 Correlation between P_{cir} and P_s in Bulk Semiconductor . . . 71
4.4 Correlation in Strained bulk semiconductor . . . 72
4.5 Conclusion . . . 78

5 Mechanism of induced birefringence in semiconductors 79
5.1 Anisotropic momentum space filling . . . 80
5.2 Birefringence within the "dephasing" time . . . 83
5.2.1 Effective two-band model . . . 83
5.2.2 Phenomenological relaxation times T_1 and T_2 . . . 85
5.2.3 Formal solution of the equations of motion . . . 85
5.2.4 Adiabatic following approximation . . . 88
5.2.5 Non-overlapping ultrashort linearly polarized pulses . . . 89
5.3 Conclusion . . . 94

6 Polarization dependent FWM quantum beats 95
6.1 Introduction . . . 95
6.2 Model . . . 96
6.2.1 Dipole matrix elements . . . 96
6.2.2 Four-wave mixing . . . 97
6.3 Conclusion . . . 106

Bibliography 108

Publications 121

Summary **123**

Samenvatting **125**

Acknowledgements **127**

Curriculum Vitae **128**

Chapter 1

Introduction

The goal of this thesis is to study some problems concerning the processing of ultrashort optical pulses and their interaction with semiconductors, which are both academically interesting and technically important. The study includes the investigation of the ultrashort optical pulse propagation and coupling dynamics in the nonlinear coupled waveguide [1], and the subpicosecond dynamics of light-semiconductor interactions [2]. The latter investigations mainly focus on polarization dependent excitations and the related electron anisotropic momentum space filling, transient birefringence of semiconductors, induced by ultrashort pulses, the correlation of the electron spin states and polarization of the emitted light in anisotropic semiconductors, the polarization dependent coherent nonlinear effects and the related items. The whole work is supported by the "Towards Ultrafast Communications (TUC)" project, with its motivation to implement 10^{12}bit/s or THz optical telecommunications in the future. This requires the generation, control and processing of subpicosecond optical pulses and the comprehensive understanding of the interaction of such pulses with semiconductors, the main materials used in optoelectronic devices.

In the already commercialized Gbit/s optical telecommunication networks, the optical signal pulse width is around the nanosecond. Therewith the pulse peak intensity is low and the whole optical system is running in the linear domain. As the data rate increases from Gbit/s to Tbit/s, which requires a decrease of the pulse length from nanosecond to subpicosecond, the nonlinear effects of the system become an important issue to reach further progress, for instance the optical pulse nonlinear propagation in optical fibers and the nonlinear directional coupler. On the other hand, when data rates are higher than 40 Gbit/s, the electronic devices can not respond at so high speed, a so-called *electronic bottleneck* will form and forbid the data rate's further increasing [3].

Therefore a transparent optical network, meaning no transfer between the light domain and electric domain, is desirable. To achieve this, lots of network functional devices should be developed to replace the previous electronic ones, for instance switching, memory, logic gate, and frequency (wavelength) conversion. Semiconductor based devices are widely recognized as prime candidates for these transparent networks. Therefore a comprehensive understanding of the interaction of subpicosecond pulses with semiconductors becomes urgent and important.

In section 1.1 a short description of light propagation in nonlinear media and a nonlinear directional coupler will be presented. The interaction of light with a semiconductor will be described in section 1.2.

1.1 Nonlinear propagation of light

Nonlinear optical phenomena are "nonlinear" in the sense that they occur when the response of a material system $\mathcal{P}$ to an applied optical field depends in a nonlinear manner upon the optical field $\mathcal{E}$ [4]. It is understandable that the response of a material system cannot grow indefinitely with increasing intensity of a light field. So as the light field intensity increases, this limitation becomes more obvious. But it also can be understood that the material system itself depends on the optical field, by which it is distorted and consequently the material response becomes nonlinear. In any case, the nonlinear response of the materials can be written as [1]

$$\mathcal{P} = \epsilon_0(\chi^{(1)}\mathcal{E} + \chi^{(2)}\mathcal{E}^2 + \chi^{(3)}\mathcal{E}^3 + \ldots). \tag{1.1}$$

Silica optical fiber, the main medium in the optical telecommunication network, is a weakly nonlinear medium. Due to the symmetrical structure, the lowest order of nonlinearity in the silica fiber is the third order $\chi^{(3)}$. A nonlinearity coefficient n_2 is often used [1]

$$n_2 = \frac{3}{8n}\mathrm{Re}(\chi^{(3)}_{\chi\chi\chi\chi}). \tag{1.2}$$

The typical nonlinearity coefficient of glass is $n_2 = 3 \times 10^{-11}\mathrm{m}^2/\mathrm{GW}$ [3].

When the data rate of the optical network is low, e.g. Gbit/s, the optical signal repetition permits a relatively wide signal pulse width. Therefore even though the total power of the optical signals is relative high, the nonlinear effects of the system will not be a problem since the peak intensity still remains in the lower level, so the third order term in Eq.(1.1) can be neglected.

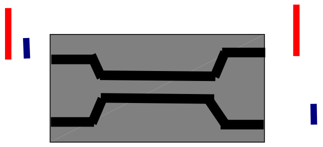

Figure 1.1: Nonlinear directional coupler: two optical signals with different intensity propagate in a NLDC and output in different channels.

When the data rate becomes high, the repetition rate of the optical signal increases; then the pulse width of the optical signal must decrease. This results in a much higher peak power of the pulse and consequently the nonlinear terms become more important.

On the other hand, the short optical pulse has a broad profile in the frequency domain. The coupling of the frequency dependent linear terms to the nonlinear terms then results in some more effects.

The light propagation in the optical fiber medium, including linear and nonlinear effects, has been documented in many textbooks [1,5]. The nonlinear Schrödinger equations (NLSEs) play the main role in the numerical investigation of light propagation in a nonlinear medium:

$$\frac{\partial A}{\partial z} = -\beta_1 \frac{\partial A}{\partial t} - i\frac{\beta_2}{2}\frac{\partial^2 A}{\partial t^2} + \frac{\beta_3}{6}\frac{\partial^3 A}{\partial t^3} - \frac{\alpha}{2}A + i\gamma|A|^2 A, \tag{1.3}$$

where A is the slowly varying envelop amplitude of the electric field of the light. The envelop moves with the group velocity $v_g = 1/\beta_1$; β_2 is the group velocity dispersion (GVD) and β_3 is the third order dispersion; γ and α the nonlinearity and loss (gain) coefficient in the optical fiber, respectively. Several methods are developed to solve this kind of equations, e.g. split-step Fourier method, and finite differential method [1].

In the optical network, the optical coupler is the most often used optical device, next to the optical fiber. Analogous to the electronic transmission, the optical coupler just plays the role of transmitting light into different branches of the network, which is a simple crossing-point in the electronic network.

There is another type of optical coupler, called optical nonlinear directional coupler (NLDC), which is more than a coupler, as it has the function to selectively output the light in a different channel. This function is very attractive, due to the possibility to implement the all-optical self-switching function. In

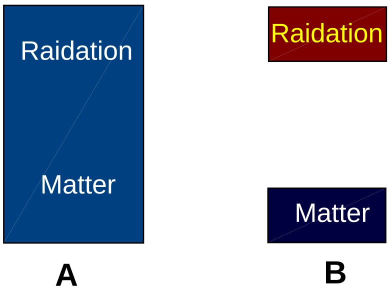

Figure 1.2: Schematic representation of the interaction between matter and light radiation. A: Quantum mechanically, matter and radiation form one complete system. B: The separate treatment of radiation and matter [6].

Fig. 1.1, a NLDC is shown, The output signal depends on the intensity of the signal itself. A switch is the simplest element of various functional devices in both the electronic and optical world. Therefore the comprehensive understanding of light propagation in a NLDC is very important for a future all-optical intelligent network.

The properties of continuous wave (cw) or quasi-cw light propagation in a NLDC are well-understood [1]. However, for ultrashort pulses, the knowledge is quite incomplete, even though much effort has been devoted to it, as analyzed in chapter 2. This thesis shows a complete description of ultrashort pulse propagation in NLDCs, and based on its switching characteristics, two new functions are proposed, rectangular optical pulse generation and optical pulse limiting.

1.2 Light-semiconductor interaction

The light-matter interaction is one of the most fundamental processes in nature. Before we discuss the more specific case, light-semiconductor interaction, we will have a look at light-atom interaction first.

The most important and simple model of light-atom interaction is a reduction of the light field to a single mode and approximation of the atom by a quantum mechanical two-level system. In this description, there is no unique distinction between matter and light: photons are shared between matter and radiation field modes, and the correlations between them make it impossible

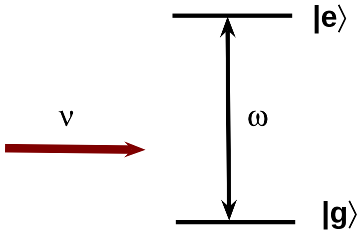

Figure 1.3: Schematic representation of the interaction between light and a two-level system.

to split the system into two parts that can be discussed independently as we schematically indicate in Fig. 1.2. However here we still adopt a semi-classical description of the light-atom interaction.

Consider the interaction of a single-mode radiation of frequency ν with a two-level atom as shown in Fig.1.3. Let $|g\rangle$ and $|e\rangle$ represent the ground and excited states of the atom system. They are eigenstates of the unperturbed part of the Hamiltonian H_0 with the system eigenvalues ϵ_g and ϵ_e, respectively.

Besides the probabilities to find the atomic *system* in the states $|g\rangle$ or $|e\rangle$, the state of the system is characterized by its momentary value of the transition dipole moment

$$\langle \mathcal{O}_e | \, \hat{r} \, | \mathcal{O}_g \rangle \, , \tag{1.4}$$

which couples directly to the electric field component of the light. The semiconductor Bloch equations (SBE), which form the main method in the literature to study the light-semiconductor interaction [2,7] use the transition dipole moment density, determined by the one-body density matrix.

When the time scale of interest is longer than the coherence time scale, we normally can talk of electrons and holes as *real* particles in ground or excited state. However in case that we want to study the coherent processes that happen in a semiconductor, the coherence correlation of the electron system will play the main role. In that case it is advantageous to cast the dynamical equations in such a form that the coherence properties are fully exploited and emphasized. This will play a role in the final chapters of this thesis.

1.3 Outline of the thesis

In chapter 2, we will describe an extended split-step Fourier method (SSFM), which will be used in the numerical calculations. Then the ultrashort optical pulse propagation in NLDCs will be studied. Based on this, the ultrashort rectangular optical pulse generation will be presented first, and then the pulse limiting feature in asymmetrical NLDCs with one self-focusing and one self-defocusing core will be demonstrated.

Chapter 3 deals with some fundamentals of semiconductors, which will be used in chapter 4-6. The electronic wave functions and energy bands are given for a semiconductor with and without strain. Thereafter we consider the optical properties of a semiconductor, focusing on the polarization dependence of optical transitions. Based on this the polarization dependent gain of a semiconductor optical amplifier (SOA) is analyzed. Next, the mechanisms of the optical signal processing by a SOA-based device and the nonlinearities of SOAs are discussed.

Chapter 4 is devoted to the study of the correlation between the electron spin polarization and circular polarization of the emitted light. This is important for the future ultrafast and nano-scale spintronics. We rigorously derive and prove a commonly used formula for the correlation in case of a bulk semiconductor. Also, we investigate the relation in the uniaxially strained bulk semiconductor. We notice that, in uniaxially strained bulk semiconductors, it is possible to achieve high spin polarization of the electrons in certain parts of the energy spectrum by circularly polarized light. We also suggest a method for detection of strain in bulk semiconductor by measuring the circular polarization of its luminescence.

Chapter 5 studies the induced transient birefringence in semiconductors due to the ultrashort pulse. We investigate the dipole polarization non-equilibrium in an optically exited semiconductor, its influence on the response to the polarized electric field, and the resulting ultrashort birefringence. This is done by the microscopic Heisenberg equations of motion, or semiconductor Bloch equations.

Chapter 6 addresses the polarization dependence of the coherent light-semiconductor interaction on an ultrashort time scale within a model that emphasizes the coherence of the photo-excited electron-hole pairs correlation. Based on this model, the polarization dependence of quantum beats in four wave mixing signal of semiconductors is studied.

Chapter 2

Manipulation of ultrashort optical pulses by nonlinear directional couplers

Nonlinear directional couplers (NLDC) are four-port devices in which an optical beam can be switched from one port to another by changing the input light power [1, 8], or pulse-width in case of incident pulses [9]. Such couplers can be made by using dual-core fibers in which the two cores are close enough, so that the evanescent wave coupling between the optical modes associated with each core transfers power from one core to another.

The nonlinear directional coupler has been extensively investigated for its potential applications in the optical signal processing and communication systems [1]. Based on its power-dependent transmission characteristics, many interesting applications have been proposed, such as all-optical switching and modulation [8,10], optical compression [11], all-optical digital switching [12,13], logic operations [14] and all-optical bistability [15]. The all-optical switching in the single NLDC and cascaded NLDCs have also been demonstrated experimentally [16–20].

In this chapter, we will demonstrate two manipulations of ultrashort optical pulses based on NLDCs: the ultrashort rectangular pulse generation and all-optical ultrashort pulse limiting. These two applications arise from the new understanding of the ultrashort optical pulse coupling dynamics in the NLDC. In the literature it had been studied in soliton format, but this did not reveal all the features of pulse coupling dynamics in an NLDC. For instance it does not provide the criteria required to qualify as a good switching [21].

In section 2.1 we will describe an extended split-step Fourier method (SSFM), which will be used in the numerical calculation. In section 2.2 the ultrashort optical pulse propagation in NLDCs will be studied. The ultrashort rectangular pulse generation will be presented in section 2.3, and in section 2.4 the pulse limiting feature in asymmetrical NLDCs will be demonstrated.

2.1 Solving coupled nonlinear Schrödinger equations[1]

The optical pulse propagation in a nonlinear optical waveguiding system can be described by the nonlinear Schrödinger equations (NLSEs) [1]. However, except in some special cases, it is not possible to solve the NLSEs analytically when both the nonlinear and dispersion (linear) effects are considered. Various numerical methods have been developed for the calculations; the most and widely used method in studying the transmission of optical pulses in a single mode fiber is the split-step Fourier method (SSFM) because of its simplicity, flexibility, good accuracy, and relatively modest computing cost [1]. In this chapter we describe a numerical method to solve the NLSEs with coupling terms, based on the SSFM method. This extended SSFM is useful to investigate the optical pulse propagation in the NLDC.

First we write the NLSE that governs the optical pulse propagation in a single mode fiber waveguide [1]

$$\frac{\partial A}{\partial z} = -\beta_1 \frac{\partial A}{\partial t} - i\frac{\beta_2}{2}\frac{\partial^2 A}{\partial t^2} + \frac{\beta_3}{6}\frac{\partial^3 A}{\partial t^3} - \frac{\alpha}{2}A + i\gamma |A|^2 A \tag{2.1}$$

where A is the slowly varying envelop amplitude of the electrical field of the light; the envelop moves at the group velocity $v_g = 1/\beta_1$, β_2 is the group velocity dispersion (GVD), and β_3 is the third order dispersion, γ and α the nonlinearity and loss (gain) coefficient in the waveguide, respectively. Here we take only instantaneous nonlinear effects into account, which is fair for the optical pulse with pulse width longer than 100 fs [1]. The influence of non-instantaneous process is studied in [22]. The normal numerical method to solve this kind of nonlinear equation is the split-step Fourier method (SSFM), which has been described in [1].

However, sometimes we have to solve N coupled NLSEs with linear coupling

---

[1] Y. Wang and W. Wang, IEEE Photon. Tech. Lett. **16**. 1077, (2004)

terms, like

$$\frac{\partial A_n}{\partial z} = -\beta_{n1}\frac{\partial A_n}{\partial t} - i\frac{\beta_{n2}}{2}\frac{\partial^2 A_n}{\partial t^2} + \frac{\beta_{n3}}{6}\frac{\partial^3 A_n}{\partial t^3} - \frac{\alpha_n}{2}A_n$$
$$+i\gamma_n|A_n|^2 A_n + \sum_{j=1,j\neq n}^{N}\left(i\kappa_{n,j}A_j - \eta_{n,j}\frac{\partial A_j}{\partial t}\right) \qquad (2.2)$$

where A_n with $n = 1, 2, 3..., N$ is the slowly varying envelop amplitude of the light field in waveguide n; β_{n1}, β_{n2} and β_{n3} are, respectively, the first, second and third order dispersion, γ_n and α_n the nonlinearity and loss (gain) coefficient in the waveguide n, respectively; κ_{nj} is the linear coupling coefficient between waveguide n and j; η_{nj} is the intermodal dispersion (IMD) coefficient [30] or the first-order coupling dispersion coefficient [31].

By taking the transformation $T = t - (1/N)\sum_{n=1}^{N}\beta_{n1}z$, Eq. (2.2) becomes

$$\frac{\partial A_n}{\partial z} = \delta_{n1}\frac{\partial A_n}{\partial T} - i\frac{\beta_{n2}}{2}\frac{\partial^2 A_n}{\partial T^2} + \frac{\beta_{n3}}{6}\frac{\partial^3 A_n}{\partial T^3} - \frac{\alpha_n}{2}A_n$$
$$+i\gamma_n|A_n|^2 A_n + \sum_{j=1,j\neq n}^{N}\left(i\kappa_{n,j}A_j - \eta_{n,j}\frac{\partial A_j}{\partial T}\right). \qquad (2.3)$$

Here $\delta_{n1} = \sum_{n=1}^{N}\beta_{n1}/N - \beta_{n1}$ is the linear mismatch.

Following the SSFM method, Eq. (2.3) can be numerically solved. In the first step, nonlinearity acts alone. In the second step, only the linear terms act. Hence, Eq. (2.3) can be spilt into a nonlinear part

$$\frac{\partial A_n}{\partial z} = i\gamma_n|A_n|^2 A_n,$$

and a linear part

$$\frac{\partial A_n}{\partial z} = \delta_{n1}\frac{\partial A_n}{\partial T} - i\frac{\beta_{n2}}{2}\frac{\partial^2 A_n}{\partial T^2} + \frac{\beta_{n3}}{6}\frac{\partial^3 A_n}{\partial T^3} -$$
$$\frac{\alpha_n}{2}A_n + \sum_{j=1,j\neq n}^{N}\left(i\kappa_{n,j}A_j - \eta_{n,j}\frac{\partial A_j}{\partial T}\right).$$

The solution of the nonlinear part can be numerically obtained by an iterative scheme

$$A_n(z+h,t) = A_n(z,t)\exp\left(i\gamma_n|A_n|^2 h\right). \qquad (2.4)$$

To solve the linear part, we work in the frequency domain,

$$\frac{\partial \widetilde{A}_n}{\partial z} = g_n \widetilde{A}_n + \sum_{j=1, j\neq n}^{N} iC_{n,j} \widetilde{A}_j \tag{2.5}$$

where $\widetilde{A}(z,\omega)$ is the Fourier transform of $A(z,t)$,

$$g_n = -i\delta_{n1}\omega + (i\omega^2\beta_{n2})/2 + (i\omega^3\beta_{n3})/6 - (\alpha_n/2) \tag{2.6}$$

and

$$C_{n,j}(z,\omega) = \kappa_{n,j} + \omega\eta_{n,j}(z). \tag{2.7}$$

Eqs. (2.5) are a set of N coupled linear first-order differential equations.

In order to solve Eqs. (2.5), we assume the solutions to take the form

$$\widetilde{A}_n(z,\omega) = f(z,\omega) \exp \int_0^z \mathrm{d}z'\ g_n. \tag{2.8}$$

Substituting this in Eqs. (2.5), we get

$$\frac{\partial f}{\partial z} = \exp\left(-\int_0^z \mathrm{d}z'\ g_n\right) \sum_{j=1, j\neq n}^{N} iC_{n,j}(z,\omega)\widetilde{A}_j(z,\omega). \tag{2.9}$$

Considering that g_n is a constant, the solution is

$$f(z,\omega) = \int_0^z \mathrm{d}z' \sum_{j=1, j\neq n}^{N} iC_{n,j}(z',\omega)\widetilde{A}_j(z',\omega)\exp(-g_n z') + D \tag{2.10}$$

where D is the integration constant. Therefore the final solution of Eqs. (2.5) is

$$\widetilde{A}_n = \exp(g_n z)\left[\int_0^z \mathrm{d}z' \sum_{j=1, j\neq n}^{N} iC_{n,j}(z',\omega)\widetilde{A}_j \exp(-g_n z') + \widetilde{A}_n(0,\omega)\right], \tag{2.11}$$

where $\widetilde{A}_n(0,\omega)$ is the input pulse at $z = 0$. By using this result, the iterative scheme of linear terms becomes

$$\begin{aligned} A_n(z+h,\omega) = \exp(g_n h)A_n(z,\omega) + \exp(g_n h)\exp(g_n z) \times \\ \left[\int_z^{z+h} \mathrm{d}z' \sum_{j=1, j\neq n}^{N} iC_{n,j}(z',\omega)\widetilde{A}_j(z',\omega)\exp(-g_n z')\right]. \end{aligned} \tag{2.12}$$

These equations can be further reduced to

$$A_n(z+h,\omega) = \exp(g_n h)\left[A_n(z,\omega) + \sum_{j=1,j\neq n}^{N} ihC_{n,j}(z,\omega)\widetilde{A}_j(z,\omega)\right]. \tag{2.13}$$

Using the iterative schemes, Eq.(2.4) and Eq.(2.13), the NLSEs with coupling terms can be numerically solved.

2.2 New normalization to study optical pulse coupling dynamics [2]

The light propagation in a NLDC has been classified into two categories: continuous wave (CW) coupling and pulse coupling. The CW or quasi-CW coupling in a NLDC obeys Jensen's equations [8]. The pulse coupling in a NLDC obeys the coupled time-dependent nonlinear Schrödinger equations (NLSEs) [1, 17, 23, 26–29], which are usually normalized in soliton fromat [26]. However this normalization can not show the full aspects of the pulse coupling dynamics. Here we will introduce a new normalization, which is more efficient than the soliton format normalization and which reveals some new features. We have found that even an ultrashort optical pulse can follow Jensen's equations which were thought to be typically applicable for the CW case only; and the switching power of a NLDC depends on the input pulse width [9].

In this section we will present a new normalization for the case of light propagation in two-core coupled waveguides. First we write the nonlinear coupled equations for a symmetrical NLDC in general form

$$\begin{aligned}\frac{\partial a_j}{\partial z} = -\beta_1\frac{\partial a_j}{\partial t} - \frac{i}{2}\beta_2\frac{\partial^2 a_j}{\partial t^2} + \frac{1}{6}\beta_3\frac{\partial^3 a_j}{\partial t^3} - \frac{\alpha}{2}a_j + \\ i\gamma|a_j|^2 a_j + i\kappa a_{3-j} - \eta\frac{\partial a_{3-j}}{\partial t}.\end{aligned} \tag{2.14}$$

Here a_j is the slowly-varying envelope amplitude of the field in channel j of the NLDC with $j = 1, 2$; β_1, β_2 and β_3 are the first, second and third order dispersion coefficients, respectively; γ, α are the nonlinearity coefficient and loss (gain) coefficient; κ is the linear coupling coefficient [1], η_{nj} is the intermodal dispersion (IMD) coefficient [30] or the first-order coupling dispersion coefficient [31].

---

[2] Y. Wang and W. Wang, Appl. Phys. B, **79**. 51, (2004)

In the case of simple step index fiber couplers, the linear coupling coefficient between two cores can be expressed as [45]

$$\kappa = \frac{1}{k\rho^2 n_1}\frac{U^2}{V^2}\frac{K_0(Ws/\rho)}{K_1^2(W)}, \tag{2.15}$$

where K_0, K_1 are the modified *Bessel* functions, ρ is radius of the fiber core and s is the distance between the center of the two cores. The profile height parameter Δ is given by

$$\Delta = (n_1^2 - n_2^2)/2n_1^2, \tag{2.16}$$

with n_1 and n_2 the refractive index of the core and cladding respectively and the normalized frequency V is introduced as

$$V = k\rho n_1\sqrt{2\Delta}, \tag{2.17}$$

where $k = 2\pi/\lambda$ is the free-space wave vector with λ the wavelength. The core and cladding parameters U and W are given by

$$\begin{aligned} U &= k\rho\sqrt{n_1^2 - (\beta/k)^2} \\ W &= \sqrt{V^2 - U^2}, \end{aligned} \tag{2.18}$$

with β the propagation constant.

The IMD coefficient is defined as

$$\eta = \frac{\mathrm{d}\kappa}{\mathrm{d}\omega}. \tag{2.19}$$

By introducing an auxiliary function

$$\begin{aligned} g = \left[2 + 2W\frac{K_0(W)}{K_1(W)} - W\frac{s}{\rho}\frac{K_1(Ws/\rho)}{K_0(Ws/\rho)}\right]\left[1 + \frac{U^2}{W^2}\cdot\frac{K_0^2(W)}{K_1^2(W)}\right] \\ -\left[1 + 2\frac{K_0^2(W)}{K_1^2(W)}\right] \end{aligned} \tag{2.20}$$

we have [21, 30, 31]

$$\eta = \frac{\lambda\kappa}{2c\pi}g. \tag{2.21}$$

The usual soliton format normalization is taken as, following [1] without taking higher order dispersion, IMD, and loss into account

$$\begin{aligned} \frac{\partial u}{\partial \xi} &= -i\frac{\mathrm{sgn}(\beta_2)}{2}\frac{\partial^2 u}{\partial \tau^2} + i|u|^2 u + Kv, \\ \frac{\partial v}{\partial \xi} &= -i\frac{\mathrm{sgn}(\beta_2)}{2}\frac{\partial^2 v}{\partial \tau^2} + i|v|^2 v + Ku. \end{aligned} \tag{2.22}$$

Here $K = \kappa L_D$, with L_D the second order dispersion length

$$L_D = \frac{T_0^2}{|\beta_2|}, \tag{2.23}$$

with T_0 the pulse width [1], $\xi = z/L_D$ the normalized length, $\tau = (t - \beta_1 z)/T_0$, and the normalized amplitudes $u = a_1\sqrt{\gamma L_D}$, $v = a_2\sqrt{\gamma L_D}$. One should notice that both normalized length ξ and amplitude unit $\sqrt{\gamma L_D}$ are dependent on the input pulse in the soliton format normalization.

The new normalization suggested by us is written as

$$\begin{aligned}\frac{\partial a_j}{\partial Z} = -i\frac{\mathrm{sgn}(\beta_2)}{2\kappa L_D}\frac{\partial^2 a_j}{\partial \tau^2} + \frac{\mathrm{sgn}(\beta_3)}{6\kappa L'_D}\frac{\partial^3 a_j}{\partial \tau^3} + i\frac{\gamma}{\kappa}|a_j|^2 a_j \\ -\frac{\alpha}{2\kappa}a_j + ia_{3-j} - \frac{\eta}{T_0\kappa}\frac{\partial a_{3-j}}{\partial \tau},\end{aligned} \tag{2.24}$$

in which the third order dispersion length is

$$L'_D = \frac{T_o^3}{|\beta_3|}, \tag{2.25}$$

Z is the normalized length, $Z = z\kappa$. Let $A_j = a_j/\sqrt{P_c}$ and $P_c = 4\kappa/\gamma$ which is the critical power [1,8]. We then get

$$\begin{aligned}\frac{\partial A_j}{\partial Z} = -i\frac{\mathrm{sgn}(\beta_2)}{2\kappa L_D}\frac{\partial^2 A_j}{\partial \tau^2} + \frac{\mathrm{sgn}(\beta_3)}{6\kappa L'_D}\frac{\partial^3 A_j}{\partial \tau^3} + i\frac{\gamma}{\kappa}|A_j|^2 A_j \\ -\frac{\alpha}{2\kappa}A_j + iA_{3-j} - \mathrm{IMD}\frac{\partial A_{3-j}}{\partial \tau}.\end{aligned} \tag{2.26}$$

with

$$\mathrm{IMD} = \eta/(T_0\kappa), \tag{2.27}$$

which depends on η, input pulse width and coupling coefficient. We note that the influence of the intermodal dispersion on the system is determined by IMD rather than η. By the new normalization, both normalized amplitude unit and normalized length Z are independent on the input pulse. Obviously, the coupling behavior of pulses depends on the parameters: $\kappa L_D = \pi L_D/2L_c$, $\kappa L'_D = \pi L'_D/2L_c$ with $L_c = \pi/(2\kappa)$ the half beat length, and IMD. Essentially, these parameters are determined by the input pulse width T_0, wavelength λ and the physical properties of the coupler waveguide.

If $\kappa L_D >> 1$, $\kappa L'_D >> 1$, $\alpha = 0$, and $|\mathrm{IMD}| << 1$, Eqs. (2.26) reduce to Jensen's equation [8]:

$$\frac{\partial a_j(\tau, Z)}{\partial Z} = i4|a_j(\tau, Z)|^2 a_j(\tau, Z) + ia_{3-j}(\tau, Z). \tag{2.28}$$

Eq. (2.28) has an exact solution in the case of single-input. It can be expressed as [8, 13]:

$$P_1(\tau, Z) = \frac{P_1(\tau, 0)}{2} \cdot \begin{cases} 1 + \mathrm{cn}\left(2Z | P_1^2(\tau, 0)\right) & \text{if } P_1(\tau, 0) < 1 \\ 1 + \mathrm{sech}(2Z) & \text{if } P_1(\tau, 0) = 1 \\ 1 + \mathrm{dn}\left(2Z P_1(\tau, 0) | \frac{1}{P_1^2(\tau, 0)}\right) & \text{if } P_1(\tau, 0) > 1 \end{cases} \tag{2.29a}$$

and

$$P_2(\tau, Z) = P_1(\tau, 0) - P_1(\tau, Z), \tag{2.29b}$$

where $P_j(\tau, Z) = |a_j(\tau, Z)|^2$ and the light is assumed to be launched into waveguide 1 only. Hence the total input power is $P_1(\tau, 0)$ and $P_2(\tau, 0) = 0$. The cn and dn in Eq. (2.29) are *Jacobi* elliptic functions. It is obvious that an input pulse will break up into two channels at the output of a NLDC depending on its instantaneous power [16, 36].

2.3 Ultrashort rectangular optical pulse generation[3]

2.3.1 Introduction

Ultrashort rectangular optical pulses are desirable for a wide range of ultrafast pump-probe experiments, studies of carrier dynamics or coherent excitation and control of optically induced quantum states [32, 33]. The use of the Fourier synthesis technique, that was first applied to produce ultrashort rectangular pulses [34], has problems as the use of the lossy and expensive, bulky, optical elements. To avoid these problems, recently the superstructured fiber Bragg gratings were used to generate a rectangular pulse, and a 20 ps rectangular pulse was demonstrated experimentally by using a 2.5 ps soliton [35]. However, in this method a well-defined input pulse shape is needed which makes the input system complex. Moreover the width of the generated rectangular pulse becomes much wider than the input pulse width. Here we demonstrate an alternative way for the generation of ultrashort rectangular pulses by using nonlinear directional couplers (NLDCs). In this new method, there are no special requirements on the input pulse shape, and the generated output pulse width is shorter than the input pulse.

Much attention has been paid to efficient ultrashort optical pulse power transfer between two coupled waveguides in a NLDC for its application in all-optical switching [20, 36]. Also there are some other applications that do not

---

[3] W. Wang, Y. Wang, K. Allaart and D. Lenstra, Opt. Commun. **253**. 164, (2005)

require complete power transfer between the two channels of a NLDC and are based on the continuous wave (CW) coupling characteristics [11, 14, 15, 37]. Recently we found that the product of dispersion length L_D and coupling coefficient κ is the key parameter of a NLDC. The pulse coupling behavior mainly depends on κL_D rather than on the input pulse shape when the third-order dispersion and intermodal dispersion are negligible [23]. Also it was found that, regarding the energy transfer, the pulse switching performance in a NLDC is similar to that of the CW case when $\kappa L_D > 10$. In a normal coupler with second order dispersion coefficient $\beta_2 = -6.5$ ps^2/km and half-beat length $L_c =$ 5mm, κL_D could be around 70, even for a 40 fs optical pulse. Therefore, from the viewpoint of energy transfer, an ultrashort pulse could exhibit the same coupling behavior in a NLDC as in the CW case. Based on this, we numerically demonstrate the generation of ultrashort rectangular pulses in a cascaded NLDC and investigate the working conditions. In the analysis, in view of the short pulse propagation, the influence of the intermodal dispersions [30] and third-order dispersions are included.

2.3.2 Theory

In the previous section, we have shown that the optical pulse propagation in coupled nonlinear optical waveguides can be described by coupled nonlinear Schrödinger equations. Here we will show the rectangular pulse generation based on a NLDC composed of two identical nonlinear optical waveguides. The equations used in the demonstration are Eqs.(2.26), the numerical method used is the iterative scheme Eq.(2.4) and Eq.(2.13).

2.3.3 Generation of a flat-top pulse

Fig. 2.1a shows the temporal profile of the output pulse from the launching channel of a NLDC with $\beta_2 > 0$. The pulse is launched into waveguide 1; that is $A_2(0, \tau) = 0$. Using Eq. (2.4) and Eq. (2.13) we assumed $\kappa L'_D = 600$, and IMD= -0.01, the normalized length of the coupler is 0.45π, $\kappa L_D = 20$ and the input pulse is $A_1(0, \tau) = 1.25 \exp(-\tau^2/2)$. Fig. 1b shows the normalized output power from the launching waveguide as a function of the input power, where $\kappa L'_D = 600$, IMD= -0.01 and the normalized length of the coupler is 0.45π. The solid line and dashed line correspond respectively to $\kappa L_D = 20$ and $\kappa L_D = 5$.

It is obvious that a Gaussian input pulse can be transformed into a flat-top pulse under certain conditions as shown as Fig. 2.1a This is due to the nonlinear restrain effect [13]. In Fig. 2.1b, it can be seen that the NLDC with

$\kappa L_D = 20$ and $\beta_2 > 0$ exhibits a nonlinear restrain effect: there is nearly no power output when the normalized input power is less than 0.4 and the output power saturates once the input power is larger than 1.1 as shown by the solid line. However, in the case of $\kappa L_D = 5$, there is no nonlinear restrain effect as shown by the dashed line. Numerical analysis shows that a flat-top pulse can be generated whenever the NLDC has $50 > \kappa L_D > 10$ in the normal dispersion regime, or $\kappa L_D > 200$ in the normal and abnormal dispersion regimes. We also investigated the influence of $\kappa L'_D$ and IMD on the output pulse shape in the NLDC. The third order dispersion term in Eqs. (2.26) may cause asymmetry and even oscillation in the falling edge of the output pulse. The influence of $\kappa L'_D$ on the output pulse shape diminishes as $\kappa L'_D$ increases. The IMD causes splitting and asymmetry of the output pulse, and its influence is reduced as the absolute value of IMD decreases. Therefore, besides the requirements on κL_D and the normalized length of coupler, to generate a symmetrical flat-top pulse the following conditions should be met at the same time: $\kappa L'_D \geq 600$ and $|\mathrm{IMD}| \leq 0.03$ when $50 > \kappa L_D > 10$, and $\kappa L'_D \geq 6000$ and $|\mathrm{IMD}| \leq 0.03$ when $\kappa L_D > 200$.

Assume a waveguide with refractive index profile $\Delta = 0.002$ and normalized frequency $V = 2.32$ at wavelength $\lambda = 0.63\mu$m, the second order and third order dispersion are $\beta_2 = 53$ ps^2/km and $\beta_3 = 0.1$ ps^3/km respectively. Let $\kappa L_D = 20$ and the half-beat length of the coupler $L_c = 1.5$cm. Then for a pulse width $T_0 = 0.1$ps we have $\kappa L'_D = 1060$, and $|\mathrm{IMD}| \approx 0.006$. They are in the range for a NLDC operating as a flat-top optical pulse generator. In fact, these operating conditions are rather relaxed, even applicable for femtosecond pulses.

2.3.4 Generation of a rectangular pulse

Clearly in Fig. 2.1 a flat-top pulse is generated with a single NLDC, but the rising and falling time of the edges are rather long compared to the pulse width. To generate a flat-top and steep-edge rectangular-like pulse, a second-order cascaded NLDC could be adopted. Fig. 2.2 shows the temporal profile of the output pulse from a second-order cascaded NLDC and its spectrum. The input pulse is $A_1(0,\tau) = 1.2\exp(-\tau^2/2)$ and $A_2(0,\tau) = 0$. The normalized coupler lengths for both couplers are $\pi/2$, $\kappa L_D = 1500$, $\kappa L'_D = 45000$ and $\mathrm{IMD} = -0.003$. Fig. 2.2a shows the output pulses from the first and the second coupler of a second-order cascaded NLDC. It is obvious that the output pulse has much steeper rising and falling edges in a second-order cascaded NLDC than in a single NLDC. Fig. 2.2b shows the spectrum of the output pulse from the launching channel of the second coupler. This spectrum is a *sinc*-like

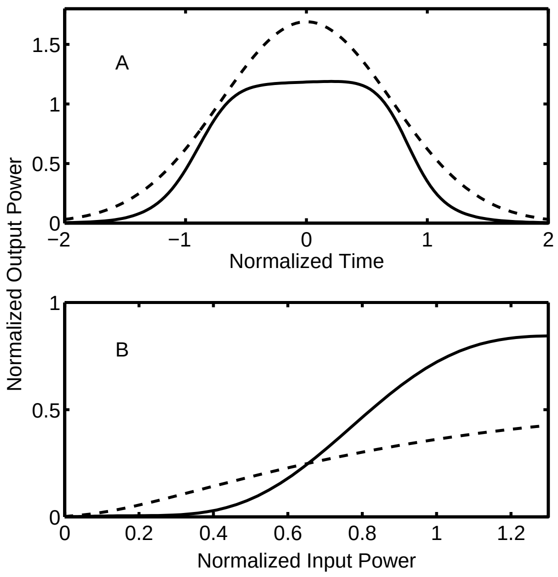

Figure 2.1: (A) Temporal profile of a flat-top pulse from the launching waveguide of a NLDC. The dotted line is the input pulse and the solid line is the output pulse. (B) The NLDC restrain effects, solid line for $\kappa L_D = 20$ and dashed line for $\kappa L_D = 5$

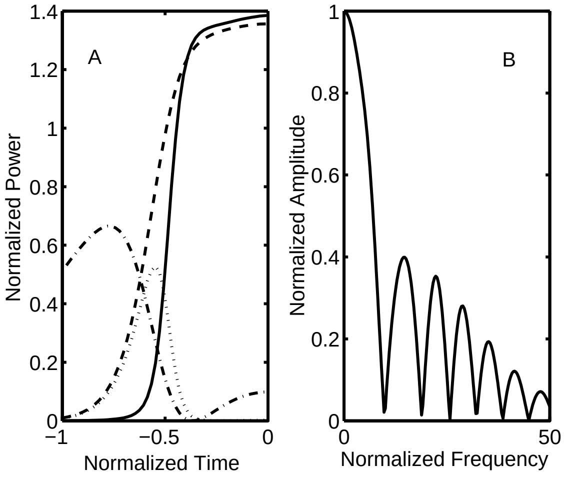

Figure 2.2: (A) Half of the symmetric temporal profile of the output pulse from a second-order cascaded NLDC, where the dashed and solid line correspond to the launching channel of the 1st coupler and 2nd coupler respectively; the dash-dotted line and dotted line correspond to the the other channel of the 1st coupler and 2nd coupler. (B) Half of the symmetric output pulse spectrum from the launching channel of the second coupler.

function.

In order to find the conditions for generating the rectangular pulse in a cascaded NLDC, we investigate the influence of κL_D, IMD, $\kappa L'_D$ and the sign of β_2 on the output pulse shape in a second-order cascaded NLDC, each coupler with a normalized length $\pi/2$. The simulation was done with the iterative scheme Eq. (2.4) and Eq. (2.13), and results are shown in Fig. 2.3, where the input is $A_1(0,\tau) = 1.2\exp(-\tau^2/2)$ and $A_2(0,\tau) = 0$. Fig. 2.3a shows the influence of κL_D on the output pulse shape from the launching channel, where the dotted line, solid line and dashed line correspond to $\kappa L_D = 2000$, $\kappa L_D = 1500$, and $\kappa L_D = 1000$ respectively; here $\kappa L'_D = 45000$, IMD$= -0.003$, and $\beta_2 < 0$ are assumed. Fig. 2.3b shows the influence of IMD on the output pulse shape from the launching channel, where the dotted line, solid line and dashed line correspond to IMD$= -0.0001$, IMD$= -0.008$, and IMD$= -0.02$ respectively. Here $\kappa L_D = 1500$, $\kappa L'_D = 5 \times 10^6$, and $\beta_2 < 0$ are assumed. Fig. 2.3c shows the influence of $\kappa L'_D$ on the output pulse shape from the launching channel, where the dotted line, solid line and dashed line correspond to $\kappa L'_D$ equal to 450000, 45000, and 4500, respectively and $\kappa L_D = 1500$, IMD$= -0.003$,

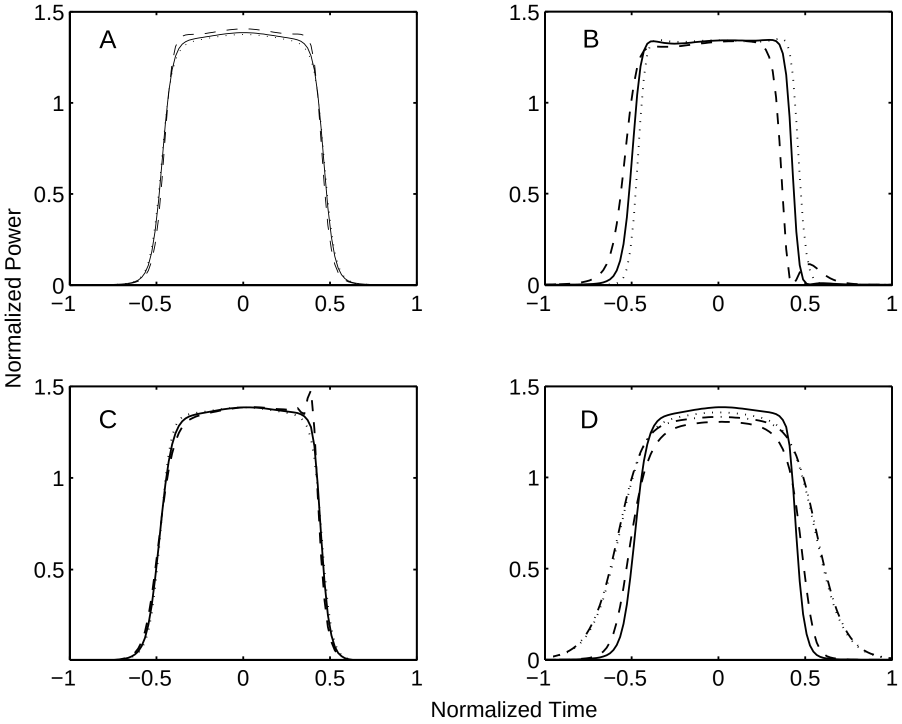

Figure 2.3: Influence on the output pulse shape in a second-order cascaded NLDC of κL_D(A), IMD(B), $\kappa L'_D$(C) and the sign of β_2(D). (A) κL_D equal to 2000(dotted), 1500(solid), and 1000(dashed); (B) IMD equal to -0.0001(dotted), -0.008(solid), and -0.02(dashed); (C) $\kappa L'_D$ equal to 450000(dotted), 45000(solid), and 4500(dashed); (D) $\beta_2 < 0$(dotted and solid) and $\beta_2 > 0$(dash-dotted and dashed).

and $\beta_2 < 0$ are assumed. Fig. 2.3d shows the influence of the sign of β_2 on the output pulse shape from the launching waveguide; the other parameters are the same as in Fig 2.2. The dotted line and solid line correspond respectively to the output of the first coupler and the second coupler in a second-order cascaded NLDC with $\beta_2 < 0$. The dash-dotted line and dashed line correspond to the output of the first coupler and of the second coupler in a second-order cascaded NLDC with $\beta_2 > 0$.

From Fig. 2.3a one can see that the pulse response of $\kappa L_D = 2000$ and of $\kappa L_D = 1500$ almost coincide, there is no ripple in the top of the pulse. When $\kappa L_D = 1000$, there is a small ripple in the top of the pulse. Therefore $\kappa L_D \geq$

1500 is required to generate a rectangular pulse in a second-order cascaded NLDC and further increasing the value of κL_D does not help much to improve the output pulse shape.

Fig. 2.3b shows the existence of a side lob in the output pulse and pulse asymmetry when IMD$= -0.02$. When IMD$= -0.008$ and IMD$= -0.0001$, there is no side lob in the output pulse. As a function of time, there is a little offset but symmetry is well preserved. Numerical analysis shows that once $|\text{IMD}| \leq 0.008$, the influence of IMD on the output pulse shape can be ignored, and the offset reduces quickly as $|\text{IMD}|$ decreases.

From Fig. 2.3c one can see that small $\kappa L'_D$ will cause asymmetry and even oscillation in the output pulse. When $\kappa L'_D = 4500$, oscillation is clearly visible in one side of the output pulse. The oscillation disappears and the pulse becomes symmetric when $\kappa L'_D \geq 45000$.

Fig. 2.3d shows the output pulse shape for $\kappa L_D = 1500$ with positive and negative β_2. The pulse shape of the output pulse of the first coupler of the second-order cascaded NLDC is not much influenced by the sign of the dispersion, like in the case of a single NLDC with $\kappa L_D \geq 200$. The output pulse of the second coupler is much better for $\beta_2 < 0$ (solid line) than for $\beta_2 > 0$ (dashed line).

From this analysis we conclude that the second-order cascaded NLDC yields a more rectangular-like output pulse shape than a single NLDC. In addition we find that much larger κL_D, $\kappa L'_D$ and smaller $|\text{IMD}|$ in a second-order cascaded NLDC do not give much improvement of the pulse shape. Therefore the conditions for generating a rectangular pulse can be summarized as: $\kappa L_D \geq 1500$, $\kappa L'_D \geq 45000$ and $|\text{IMD}| \leq 0.008$. As known, these parameters depend on the NLDC physical parameters: β_2, β_3, $|\text{IMD}|$, κ, and also on the pulse width. Since the pulse width is an important item for a pulse generator, an interesting question is, how short a pulse can be obtained in this second-order cascaded NLDC. To generate a rectangular pulse in a NLDC, the minimum input pulse width should meet the following inequalities:
for κL_D:

$$T_0 \geq \left(\frac{2|\beta_2| L_c}{\pi} \cdot 1500 \right)^{1/2}, \tag{2.30}$$

for $\kappa L'_D$:

$$T_0 \geq \left(\frac{2|\beta_3| L_c}{\pi} \cdot 45000 \right)^{1/3}, \tag{2.31}$$

and for *IMD*:

$$T_0 \geq \frac{\lambda |g|}{2\pi c \cdot 0.008}. \tag{2.32}$$

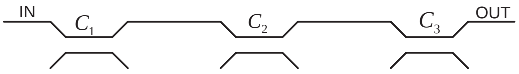

Figure 2.4: The scheme of a third-order cascaded NLDCs, with C_1, C_2 and C_3 the first, second and third coupler.

In a case with $\beta_2 = -6.5\text{ps}^2/\text{km}$, $\beta_3 = 0.1\text{ps}^3/\text{km}$ at $\lambda = 1.3\mu\text{m}$, and $L_c = 5\text{mm}$, $g \approx -3.8$ with refractive index profile: $\Delta = 0.003$, and normalized frequency $V = 2$, these three conditions imply that T_0 must be larger than 176 fs, 306 fs, and 327 fs respectively. Therefore 330 fs is the minimum working pulse width. The output pulse is significantly compressed so that about 180 fs rectangular-like pulses can be generated in this second-order cascaded NLDC.

Take the parameters γ=0.01/(W·m) [1], the critical power $P_c = 124\text{kW}$. With a general fiber core area 50∼80μm^2, this intensity is still far away from the silica fiber damage threshold [41]. Switching of a pulse with power as high as 282kW has been demonstrated in NLDC [36]. We did not include other nonlinear effects such as self-steepening and nonlinear retardation since the pulse width is larger than 100 fs [1]. It is possible to generate a shorter pulse if we decrease the half-beat length, but it implies a higher critical power. The output pulse shapes of the second-order cascaded NLDC suggest that a higher-order cascaded NLDC could generate an ideal rectangular pulse by cutting the edge much steeper. Our numerical analysis confirms this expectation but the pulse width increases correspondingly, that is, better pulse waveform is at expense of a larger pulse width. Fig. 2.4 shows the scheme of a third-order cascaded NLDC: the output of the preceding coupler is the input of the next coupler. The parameters for the three couplers C_1, C_2, and C_3 may be equal or different.

In Fig. 2.5 we show the output pulse temporal profile of C_1, C_2, and C_3. The input is $A_1(0, \tau) = 1.2 \exp(-\tau^2/2)$ and $A_2(0, \tau) = 0$. The parameters for the cascaded NLDC are: $\kappa L_D = 10^5$, $\kappa L'_D = 1.2 \times 10^6$, IMD$= -0.003$, and the normalized coupler length is $\pi/2$. The couplers C_1, C_2, and C_3 have the same parameters. With these values for C_3 only and C_1 and C_2 parameters as before in the second-order cascaded NLDC it is not possible to achieve as good results as in Fig. 2.5. The dashed line, dash-dotted line and solid line are the C_1, C_2, and C_3 output pulse shapes, respectively, and the dotted line is the input Gaussian pulse. Clearly from the C_1 to the C_3 output, the pulse shape is approaching an ideal rectangularity.

It is also possible to generate rectangular pulses with a smaller value of

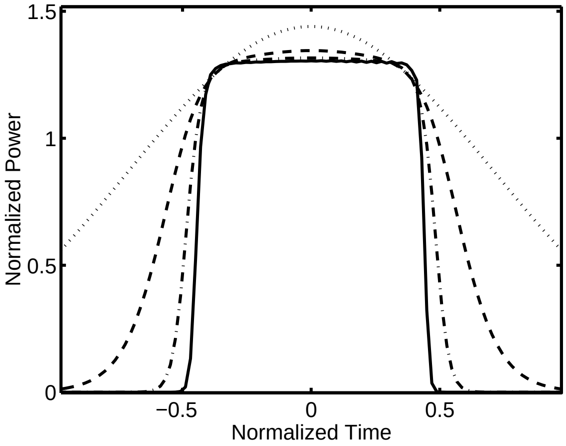

Figure 2.5: The output pulse temporal profile from the launching channel of a third-order cascaded NLDC, where the dashed, dash-dotted and solid line are the C_1, C_2, and C_3 output pulse shapes, respectively; the dotted line is the input Gaussian pulse.

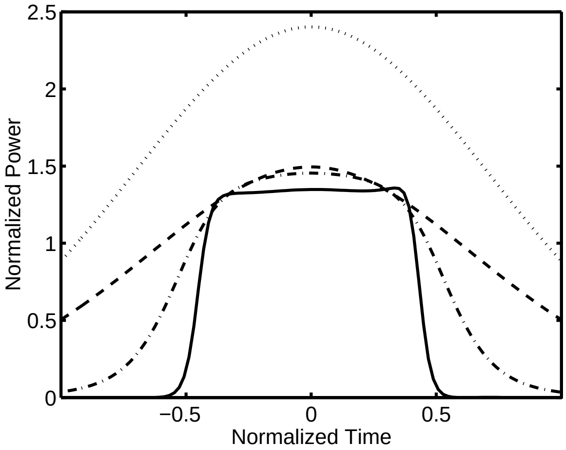

Figure 2.6: The output pulse temporal profile from the launching waveguide of a third-order cascaded NLDC with small κL_D to achieve an ultrashort rectangular pulse with moderate required pulse peak power. Dashed, dash-dotted and solid line are the C_1, C_2, and C_3 output pulse shapes; the dotted line is the input Gaussian pulse.

κL_D. We show another case with $\kappa L_D = 8000$, $\kappa L'_D = 10^6$, and IMD$= -0.001$ for C_1 and with $\kappa L_D = 2000$, $\kappa L'_D = 2.5 \times 10^5$ and IMD$= -0.001$ for C_2 and C_3. Different from the former case in which the three couplers had the same normalized length equal to $\pi/2$, we now adopt the normalized coupler length $\pi/4$ for C_1 and $\pi/2$ for C_2 and C_3. The input is $A_1(0,\tau) = 1.55\exp(-\tau^2/2)$ and $A_2(0,\tau) = 0$. The results are shown in Fig. 2.6. The dashed line, dash-dotted line and solid line are the C_1, C_2, and C_3 output pulse shape, respectively. The dotted line is the input Gaussian pulse. Obviously the peak power of the output pulse is much lower than that of the input pulse. Compared to Fig. 2.5, the output peak powers are almost equal in these two cases, but the normalized amplitude of the input pulses are different: $A = 1.2$ for the former case and $A = 1.55$ for the latter case, so the energy transfer efficiency is lower for the latter case.

We can achieve an ideal rectangular pulse in a third-order cascaded NLDC compared to the pulse generated by a second-order cascaded NLDC, but the pulse width is larger. Let us take Fig. 2.5 as an example, assuming $L_c = 5$mm, $\beta_2 = -6.5$ps^2/km, $\beta_3 = 0.1$ps^3/km, $\lambda = 1.3\mu$m, and $\Delta = 0.003$, we find that the input pulse width of this cascaded NLDC will be 1.44 ps and the output pulse width 700fs. So at the expense of the pulse width, an ideal rectangular pulse was achieved in the third-order cascaded NLDC.

2.3.5 Conclusion

We have shown femtosecond rectangular optical pulse generation for the first time by a cascade of two or three NLDCs. We presented the working conditions that should be satisfied in terms of half-beat length, second-order, third-order dispersion coefficient and coupling coefficient dispersion of the NLDC. Under realistic conditions, a 180 fs rectangular-like pulse can be generated with 5mm coupling length, and a shorter pulse could be achieved by decreasing the coupling length. In a third-order cascaded NLDC, an ideal rectangular pulse can be implemented, but the pulse width will be broader.

2.4 All-optical limiting [4]

2.4.1 Introduction

Non-linear directional couplers (NLDCs) composed of two cores with linear or nonlinear mismatch have been theoretically studied by using the coupled-

---

[4] W. Wang, Y. Wang, K. Allaart and D. Lenstra, Appl. Phys. B, **83**. 623, (2006)

mode approaches and the beam propagation method in the case of CW [24, 39, 40, 44, 46–48]. The studies so far have, however, paid little attention to the limiting characteristics of the NLDC and all of these studies focused on the CW case only. To our knowledge, there is no report yet about the optical pulse limiter based on the asymmetric NLDC composed of one self-focusing waveguide and one self-defocusing waveguide. Other passive optical limiters have been demonstrated by using optical nonlinearities, e.g. self-focusing or self-defocusing, nonlinear absorption and nonlinear scattering. These are reviewed in [42].

For an ideal optical limiter, its output should equal the input when the input power is lower than the limiting threshold, and the output is constant when the input power exceeds the limiting threshold. When working under the pulse condition, the output pulse profile should be the same as the input pulse except that its top part, which is beyond the limiting threshold, is tailored. Here we will show that an optical limiter based on linear mismatched asymmetric NLDCs composed of one self-focusing waveguide and one self-defocusing waveguide can exhibit the ideal limiting features.

It is well known that non-rectangular pulses break up when they propagate in a conventional NLDC [36], and it was demonstrated recently that pulse breakup occurs in a conventional NLDC when the product of the dispersion length and the coupling coefficient is larger than 50 [23]. However, in asymmetrical NLDCs, the rising and falling edges of the output pulse are almost the same as the input pulse. There is almost no pulse breakup. The limiting features in conventional NLDCs, whose output pulse generally suffers breakup, were used as a rectangular pulse generator [47] rather than as an optical pulse limiter. The length of asymmetric NLDCs, used in limiting, is about one beat length instead of half a beat length in the conventional NLDC, to reduce the pulse breakup.

The cascaded asymmetric NLDCs are discussed here to improve the limiting characteristics. The high order dispersion and intermodal dispersion [30, 31, 47] are included in the analysis of pulse propagation in the asymmetric NLDC.

As known, the common silica optical fiber is a self-focusing nonlinear medium, while some semiconductor doped fibers are self-defocusing nonlinear media and have large nonlinearity coefficients [43]. Since self-focusing and self-defocusing glass can be fabricated by appropriate doping, it is possible to fabricate an asymmetric coupler composed of one self-focusing waveguide and one self defocusing waveguide.

2.4.2 Theory

Coupled mode theory has proved to be a simple and reliable means to describe the coupling behavior of a NLDC with well-separated cores and constant profile for weak nonlinear perturbations [1,8,25]. By this means, the coupled equations for an asymmetric NLDC without loss can be directly written as

$$\begin{aligned}\frac{\partial a_1}{\partial z} = i\beta_{01}a_1 - \beta_{11}\frac{\partial a_1}{\partial t} - i\frac{\beta_{21}}{2}\frac{\partial^2 a_1}{\partial t^2} + \frac{\beta_{31}}{6}\frac{\partial^3 a_1}{\partial t^3} + \\ i\gamma_1|a_1|^2 a_1 + i\kappa a_2 - \eta\frac{\partial a_2}{\partial t}\end{aligned} \tag{2.33a}$$

$$\begin{aligned}\frac{\partial a_2}{\partial z} = i\beta_{02}a_2 - \beta_{12}\frac{\partial a_2}{\partial t} - i\frac{\beta_{22}}{2}\frac{\partial^2 a_2}{\partial t^2} + \frac{\beta_{32}}{6}\frac{\partial^3 a_2}{\partial t^3} - \\ i\gamma_2|a_2|^2 a_2 + i\kappa a_1 - \eta\frac{\partial a_1}{\partial t}\end{aligned} \tag{2.33b}$$

where $a_1(z,t)$ and $a_2(z,t)$ are the complex amplitudes of the field in waveguide 1 and 2; β_{01} and β_{02} are respectively the propagation constants of waveguide 1 and 2; β_{11} and β_{12}, β_{21} and β_{22}, β_{31} and β_{32} are respectively the 1st order, 2nd order and 3rd order dispersion of waveguide 1 and 2; γ_1 and γ_2 are the nonlinearity coefficients of waveguide 1 and 2; κ is the linear coupling coefficient [8, 23] and η is the intermodal dispersion (IMD).

In the case of $\beta_{01} = \beta_{02}$, $\beta_{11} = \beta_{12}$, $\beta_{21} = \beta_{22}$, $\beta_{31} = \beta_{32}$ and $\gamma_2 = -\gamma_1$, Eqs. (2.33) reduce to the well known coupled nonlinear Schrödinger equations [47]. At the same time, if $\kappa = 0$, hence $\eta = 0$, Eqs. (2.33) further reduce to the nonlinear Schroedinger equation, which has been extensively used to study the pulse propagation in optical fiber [1]. Eqs. (2.33) are valid if the input pulse width is larger than 1 ps [1].

Let $A_j = a_j/\sqrt{P_c}\exp(i\beta z)$ with $P_c = 4\kappa/\gamma_1$, which is the critical power [8, 23], and $\beta = (\beta_{01} + \beta_{02})/2$. Eqs. (2.33) can be normalized in the soliton format. Here we adopted a new normalization format [23]. Assume $\beta_{11} = \beta_{12} = \beta_1$, $\beta_{21} = \beta_{22} = \beta_2$, $\beta_{31} = \beta_{32} = \beta_3$ and $x = \gamma_2/\gamma_1$, apply the transformations $\tau = (t - \beta_1 z)/T_0$ and normalize coordinate $Z = z\kappa$, then Eqs. (2.33) become

$$\begin{aligned}\frac{\partial A_1}{\partial Z} = i\delta A_1 - i\frac{\beta_2}{2\kappa L_D}\frac{\partial^2 A_1}{\partial \tau^2} + \frac{\beta_3}{6\kappa L'_D}\frac{\partial^3 A_1}{\partial \tau^3} + i4|A_1|^2 A_1 + \\ iA_2 - \mathrm{IMD}\frac{\partial A_2}{\partial \tau}\end{aligned} \tag{2.34a}$$

$$\begin{aligned}\frac{\partial A_2}{\partial Z} = -i\delta A_2 - i\frac{\beta_2}{2\kappa L_D}\frac{\partial^2 A_2}{\partial \tau^2} + \frac{\beta_3}{6\kappa L'_D}\frac{\partial^3 A_2}{\partial \tau^3} - i4x|A_2|^2 A_2 + \\ iA_1 - \mathrm{IMD}\frac{\partial A_1}{\partial \tau},\end{aligned} \tag{2.34b}$$

where $\delta = (\beta_{01} - \beta_{02})/(2\kappa)$ is the normalized propagation constant difference; $L_D = T_0^2/|\beta_2|$ and $L'_D = T_0^3/|\beta_3|$ are the second and third order dispersion length respectively.

By the extended split-step Fourier method (SSFM) as proposed in [25], which is the combination of Split-Step Fourier method and integration method of differential equations, one can solve Eqs. (2.34) numerically.

2.4.3 CW analytical solution

When $\kappa L_D \gg 1$, $\kappa L'_D \gg 1$ and $|\mathrm{IMD}| \ll 1$, Eqs. 2.34 will reduce to the CW case, that is

$$\frac{\partial A_1}{\partial Z} = i\delta A_1 + i4|A_1|^2 A_1 + iA_2 \tag{2.35a}$$

$$\frac{\partial A_2}{\partial Z} = -i\delta A_2 - i4x|A_2|^2 A_2 + iA_1. \tag{2.35b}$$

In the case of $x = -1$ and $\delta = 0$, Eqs. (2.35) reduce to the well known Jensen's equations. If the nonlinear terms are ignored also, Eqs. (2.35) reduce to the linear coupled equations.

When $x = 1$, Eqs. (2.35) can be analytically solved as follows. We use the substitutions

$$A_j = B_j \exp(i\phi_j), \tag{2.36}$$

where B_j and ϕ_j ($j = 1,\ 2$) are functions of Z; ϕ_1 and ϕ_2 are the nonlinear phases of beams propagating in the self-focusing and self-defocusing waveguides respectively. Then we substitute Eq. 2.36 into Eq. 2.35 and split the real and imaginary part. The normalized power in waveguide 1 and 2 will be

$$P_j(Z) = B_j^2(Z) \qquad j = 1,\ 2. \tag{2.37}$$

Therefore the total power is

$$P_t = P_1(Z) + P_2(Z). \tag{2.38}$$

The initial phase difference is

$$\varphi = \phi_1(0) - \phi_2(0). \tag{2.39}$$

Define a constant Γ

$$\Gamma = [P_1(0)P_2(0)]^{1/2}\cos\varphi + \delta P_1(0) + P_1^2(0) - P_2^2(0). \tag{2.40}$$

The power evolution in waveguide 1 then obeys the equation:

$$\begin{aligned}\left(\frac{\mathrm{d}P_1}{\mathrm{d}Z}\right)^2 = -4\big[1+(\delta+2P_t)^2\big]P_1^2 + 4P_t\cdot\big[P_t + 2(\Gamma+P_t^2)\cdot \\ (\delta+2P_t^2)\big] - 4(\Gamma+P_t^2)^2.\end{aligned} \tag{2.41}$$

Eq. 2.41 is integrated analytically; we then get the power evolution in the two wave guides

$$\begin{aligned} P_1(Z) &= \frac{G}{2F} + \frac{[G^2-4HF]^{1/2}}{2F}\sin[\alpha_0 + \mathrm{sgn}(\sin\varphi)2\sqrt{F}Z] && \text{(2.42a)}\\ P_2(Z) &= P_t - P_1(Z) && \text{(2.42b)}\end{aligned}$$

with

$$\begin{aligned} F &= 1+(\delta+2P_t)^2\\ H &= (\Gamma+P_t^2)^2\\ G &= P_t + 2(\Gamma+P_t^2)(\delta+2P_t)\\ \alpha_0 &= \sin^{-1}\frac{2FP_t(0)-G}{\sqrt{G^2-4HF}}\end{aligned}$$

In the case of single-input excitation, Eqs. 2.42 can be further reduced.

Case 1: the power is initially launched into the self-defocusing waveguide only, $P_1(0) = 0$ and $P_2(0) = P_t$. We then have $F = 1 + (\delta + 2P_t)^2$, $G = P_t$, $H = 0$ and $\alpha_0 = -\pi/2$ and Eqs. 2.42 reduces to

$$\begin{aligned} P_1(Z) &= \frac{P_t}{1+(\delta+2P_t)^2}\sin^2\left\{[1+(\delta+2P_t)^2]^{1/2}Z\right\} && \text{(2.43a)}\\ P_2(Z) &= P_t - \frac{P_t}{1+(\delta+2P_t)^2}\sin^2\left\{[1+(\delta+2P_t)^2]^{1/2}Z\right\}. && \text{(2.43b)}\end{aligned}$$

In the case of $\delta = -0.3$, $Z = \pi$ and $P_t \leq 0.24$, the detuning between the two wave guides is sufficiently weak; nearly all the input power remains in the launching waveguide. Therefore we get $P_1(Z) \simeq 0$ and $P_2(Z) \simeq P_t$. When the input power increases, the detuning between the two waveguides becomes sufficiently high, and part of the input power will couple into the neighboring waveguide. The power coupling into the neighboring waveguide increases with the input power, and the output power of the launching waveguide shows saturation features. Therefore, the asymmetric NLDC can be used as an optical limiter, and it also means that in the pulse case the breakup of the output pulse will be reduced significantly.

Case 2: the power is initially launched into the self-focusing waveguide, $P_1(0) = P_t$ and $P_2(0) = 0$. In this case, the expressions of power evolution in the waveguides 1 and 2 are directly obtained by exchanging the subscripts of P_1 and P_2 in Eqs. (2.43). This implies that the power evolution in the asymmetric NLDC remains unchanged whenever the power is initially launched into the self-focusing waveguide or the self-defocusing waveguide. This feature is quite different from the mismatched NLDC composed of two self-focusing waveguides or two self-defocusing waveguides, where non-reciprocity is observed [46]. It should be noted that the reciprocity and Eqs. (2.43) are valid only when $x = 1$, that is when the non-linearity is equal for both wave guides.

Eqs. (2.43) also show that the maximum power transfer between the two waveguides is $P_t/\left(1 + (\delta + 2P_t)^2\right)$. When $|\delta + 2P_t| \gg 1$, most of the power remains in the initially launched waveguide; no power is transferred between the two waveguides. In the case of $P_t \ll 1$, Eqs. 2.43 reduce to the well-known solution of the linear coupler. If the linear mismatch δ is sufficiently large, most of the power will remain in the initially launched waveguide, independent of the P_t value.

2.4.4 Numerical analysis

We now numerically investigate the working conditions of an asymmetric NLDC to function as an optical limiter. In Fig. 2.7, the normalized output power as a function of the input power in the asymmetric NLDC is shown, where the solid and dashed lines correspond to linear mismatch $\delta = -2.5$ and $\delta = 2.5$; the normalized coupler length is 0.9π. In the conventional NLDC, the length of the coupler is usually a half beat length. To avoid pulse breakup, the length of the asymmetric NLDC is chosen about one beat length. Hence, nearly all the input power remains in the launching waveguide when the input power is small. In the case of $\delta = -2.5$, namely the propagation constant of the self-focusing waveguide is smaller than that of the self-defocusing waveguide, the output of the launching waveguide is linearly proportional to the input as long as the input is below the limiting threshold and there is nearly no power transfer between the two coupled waveguides. When the input exceeds the threshold power, the output of the launching waveguide becomes nearly constant over a range of power, as shown by the solid line in Fig. 2.7. In the case of $\delta = 2.5$, however,

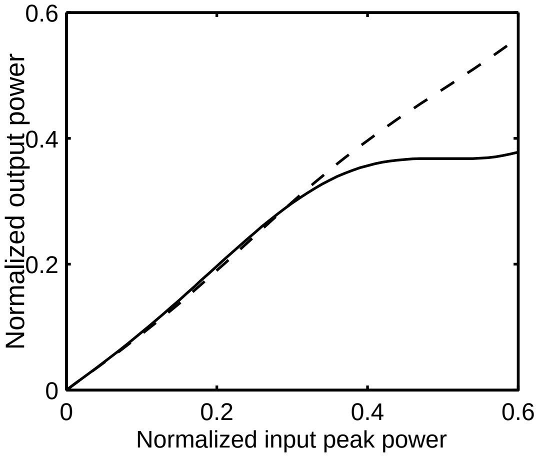

Figure 2.7: Input-output characteristics of a single-input asymmetric NLDC. Solid and dashed lines correspond to propagation constant mismatch $\delta = -2.5$ and $\delta = 2.5$.

the output always shows some linear relationship with the input. Numerical analysis shows that no limiting feature is obtained when $\delta > 0.3$ in the single-input NLDC. The obvious limiting feature can be observed when $-2.5 < \delta \leq 0$ with appropriate coupling length. Therefore, to obtain the limiting property, the propagation constant of the self-focusing waveguide should not be larger than that of the self-defocusing waveguide.

It can also be seen that both the limiting threshold and output power, normalized to P_c, are smaller than 1, unlike in the conventional NLDC, where both the limiting threshold and output are larger than 1. When the input power is lower than the limiting threshold, the output almost equals the input. As a result, the pulse breakup is reduced significantly.

It was always known that the nonlinear restraining effects of a NLDC can improve the switching characteristics by cascading several couplers together [13]. From Fig. 2.7, one can see obvious nonlinear restraining effects when the input is higher than the threshold power and a linear transmission system when the input is below the threshold. Therefore, it is possible to improve the limiting feature by cascading several couplers together.

Fig. 2.8 shows the limiting characteristics in a cascaded coupler structure. The dashed, solid and dash-dotted lines correspond to the single, second-order and third-order cascaded NLDC output. The structure of the cascaded NLDC is also shown in Fig. 2.8 where the preceding coupler output is the input of the

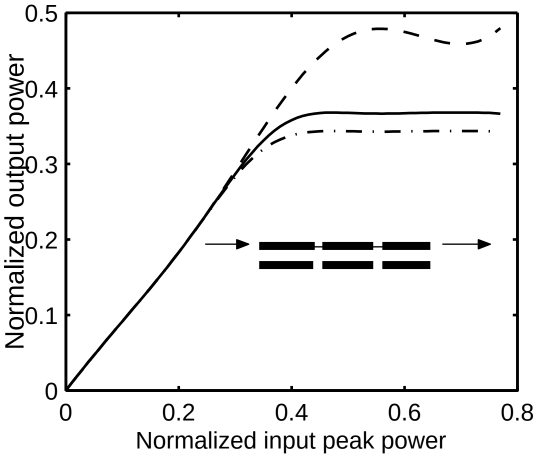

Figure 2.8: The limiting characteristics in the cascaded structure. Dashed, solid and dash-dotted lines correspond to the 1st, 2nd and 3rd coupler output.

next coupler. The physical parameters of the three couplers are assumed to be identical but with different normalized coupler lengths 0.9π, π and 1.1π, respectively. Numerical analysis shows that if the input power of the second coupler, i.e. the output power of the first coupler, is below the limiting threshold power of the second coupler, there is no improvement to the limiting characteristics, since the second coupler operates just as a linear transmission system. However, when the output power of the first coupler is above the threshold power of the second coupler, the second coupler output becomes flatter than that of the first coupler due to the nonlinear restraining effects. If the output is still not sufficiently flat, a third-order cascade NLDC can be applied to improve it further.

Application of a cascaded structure, besides flattening the output, also enables to control the output power. Fig. 2.9 shows how the length of the third coupler influences the limiting characteristics in a third-order cascaded NLDC. The normalized lengths for the first and second coupler are 0.9π and π; the linear mismatch for each coupler is $\delta = -0.3$; the solid and dashed lines correspond to the third coupler with normalized length π and 1.1π. Obviously the limiting output power decreases as the third coupler length increases, but without degrading the limiting characteristics. This provides a simple but efficient way to control the limiting output power.

It is also interesting to know the pulse response of the asymmetric NLDC, operating as a limiter. As known, pulses propagating in a conventional NLDC

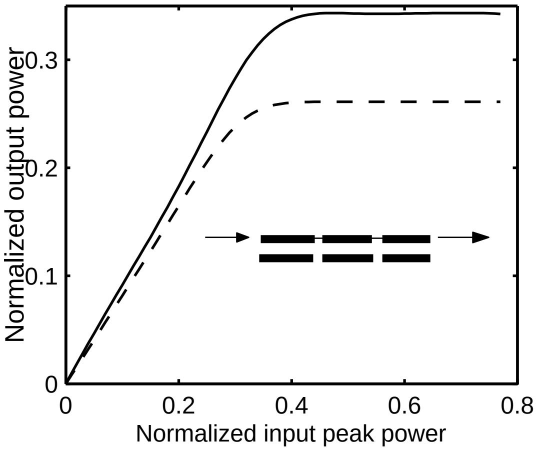

Figure 2.9: The influence of the normalized length of the third coupler on the limiting characteristics. Solid and dashed lines correspond to the third coupler normalized length $L_3 = \pi$ and 1.1π.

usually break up into two output ports according to their instantaneous power. It has been shown that the pulse breakup in a conventional NLDC will occur if the product of the coupling coefficient and the dispersion length is sufficiently large [23]. This then seriously degrades the performance of the NLDC as a switch [20, 36]. If breakup also happens in the asymmetrical NLDC, it would deteriorate the pulse response of the limiter.

Fig. 2.10 shows that the pulse response of a third-order cascaded NLDC. We assumed $\kappa L_D = 1000$, $\kappa L'_D = 150000$, $\beta_2 > 0$ and IMD $= 0$. The input pulse is $A_1(0,\tau) = u\exp(-\tau^2/2)$ and $A_2(0,\tau) = 0$ with u the normalized amplitude; in Fig. 2.10a and 2.10b, u equals $\sqrt{0.7}$ and $\sqrt{0.5}$ respectively. The dotted lines are the input pulses, the solid lines correspond to the outputs of the three couplers in the third-order cascaded NLDC with normalized coupler lengths 0.9π, π and 1.1π. Fig. 2.10 shows that, as long as the input pulse power is below the limiting threshold¡the output pulse will keep the same profile as the input pulse, so one has a linear transmission system. Where the input pulse peak power exceeds the limiting thresholds of these three couplers, the output pulse top will be tailored and the peak becomes flat, as the solid line 1, 2 and 3 in Fig. 2.10, but the rising and falling edges of the output pulse remain unchanged. In both cases, no pulse breakup occurs in this asymmetrical NLDC; even it is operated in the quasi-CW conditions. This is quite different from the pulse propagation in the conventional NLDC.

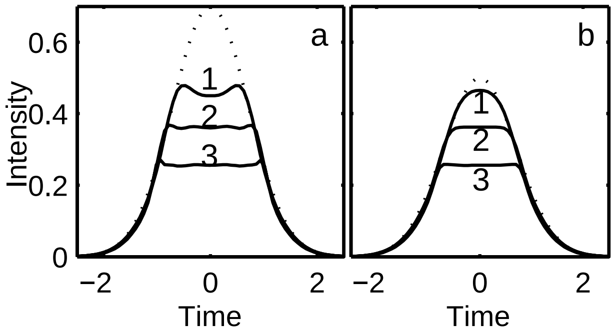

Figure 2.10: Temporal profiles of the input-output pulses. Dotted line is the input pulse profile. Solid line 1, 2 and 3 are the output of the 1st, 2nd and 3rd coupler. a. input $A_1(0,\tau) = \sqrt{0.7}\exp(-\tau^2/2)$ and $A_2(0,\tau) = 0$, b. input $A_1(0,\tau) = \sqrt{0.5}\exp(-\tau^2/2)$ and $b_2(0,\tau) = 0$.

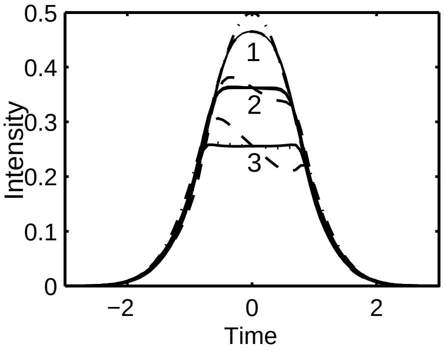

Figure 2.11: The influence of the intermodal dispersion on the pulse response of the cascaded asymmetric NLDCs. Solid, dotted and dashed lines correspond to IMD $= 0, -0.001$ and -0.01. The dash-dotted line is the input pulse. The curve 1, 2, and 3 are the outputs of the 1st, 2nd and 3rd coupler.

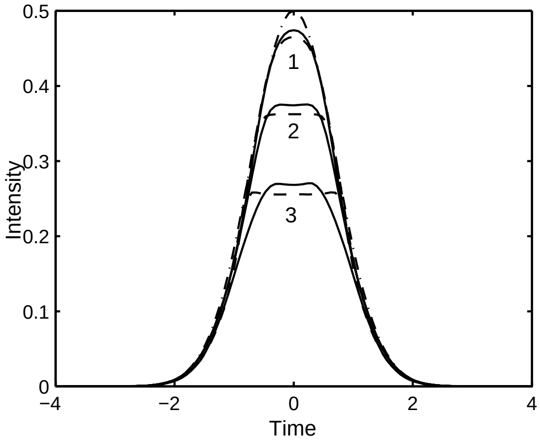

Figure 2.12: The influence of the sign of β_2 on the pulse response of the cascaded asymmetric NLDCs. Dashed and solid lines correspond to $\beta_2 > 0$ and $\beta_2 < 0$. The curve 1, 2, and 3 are the outputs of the 1st, 2nd and 3rd coupler.

Fig. 2.11 shows the influence of the intermodal dispersion on the pulse response of the cascaded asymmetric NLDC where the inputs are $A_1(0,\tau) = \sqrt{0.7}\exp(-\tau^2/2)$ and $A_2(0,\tau) = 0$. Except for IMD, the other parameters are the same as in Fig. 2.10. The solid, dotted and dashed lines correspond to IMD values equal to 0, -0.001 and -0.01. The dash-dotted line is the input pulse. Clearly IMD causes asymmetry in the top part of the output pulse. This deformation degrades the limiting features. It also means that even when the second order and higher order dispersion can be ignored, the IMD may still have to be considered. When $|\text{IMD}| < 0.001$, the influence of IMD on the pulse response is negligible.

Fig. 2.12 shows the influence of the sign of β_2 on the pulse response of the cascaded asymmetric NLDC with the same inputs as in Fig. 2.11. The following parameters are adopted: $L_1 = 0.9\pi$, $L_2 = \pi$, $L_3 = 1.1\pi$, $\delta = -0.3$, $\kappa L_D = 1000$, $\kappa L'_D = 150000$, and IMD $= 0$. The dashed and dotted lines correspond to $\beta_2 > 0$ and $\beta_2 < 0$; the dash-dotted line is the input pulse. One observes the limiting feature in both cases: $\beta_2 > 0$ and $\beta_2 < 0$. This is also expected from Eqs. (2.34): when $\kappa L_D \gg 1$, the κL_D term can be ignored. Consequently the sign of β_2 has little impact on the limiting characteristics. Comparison of the dashed line and solid line shows that better response could be obtained in the case of $\beta_2 > 0$ than in the case of $\beta_2 < 0$; the pulse breakup is improved a lot when $\beta_2 > 0$.

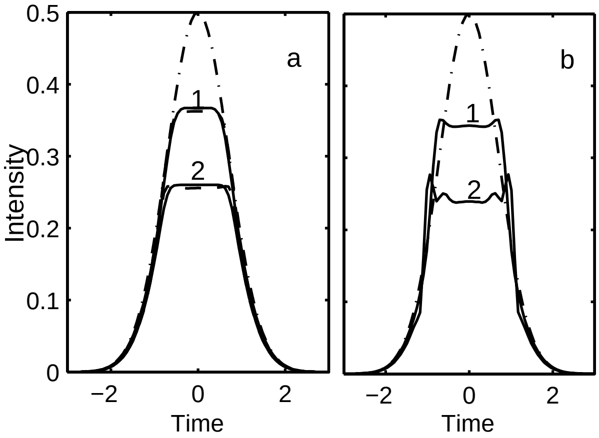

Figure 2.13: The influence of κL_D on the pulse response of the cascaded asymmetric NLDCs where curve 1 and 2 correspond to the 2nd and 3rd coupler output. Dashed and solid lines in (a) correspond to $\kappa L_D = 5000$ and 1000. The solid lines in (b) correspond to $\kappa L_D = 200$

Fig. 2.13 shows the influence of κL_D on the pulse response of the cascaded asymmetric NLDC. Except for κL_D, the other parameters are the same as in Fig. 2.12. The dashed and solid lines in Fig. 2.13a correspond to κL_D equal to 5000 and 1000; the solid line in Fig. 2.13b corresponds to κL_D equal to 200; the dash-dotted lines are the input pulses. When $\kappa L_D \gg 1000$, the influence of the κL_D on the limiting pulse response is nearly invisible as shown in Fig. 2.13a; but considerable deformation in the output pulse can be seen in Fig. 2.13b with $\kappa L_D = 200$. Therefore $\kappa L_D \geq 1000$ is required to generate a flat-top pulse in a third-order cascaded NLDC.

Fig. 2.14 shows the influence of the ratio x of the nonlinearity coefficients of the self-defocusing and self-focusing waveguide on the pulse response of the cascaded asymmetric NLDCs, where we assumed $L_1 = 0.9\pi$, $L_2 = \pi$, $L_3 = 1.1\pi$, $\delta = -0.3$, $\beta > 0$, $\kappa L_D = 1000$ and IMD $= 0$. The input is $b_1(0,\tau) = \sqrt{0.7}\exp(-\tau^2/2)$ and $b_2(0,\tau) = 0$. The limiting threshold and the output power increase with the reduction of x. The influence of x on the limiting performance, such as pulse breakup and the fluctuation of the output, is small as shown in Fig. 2.14.

For a case with realistic parameters: $\beta_2 = -6.5\text{ps}^2/\text{km}$, β_3=$0.1\text{ps}^3/\text{km}$, $\lambda = 1.55\mu\text{m}$, half-beat length 5mm and a pulse width of 1 ps, we have $\kappa L_D = 48000 >> 1000$, $\kappa L'_D = 3140000$ and IMD $= 0.0001$. Therefore the condition,

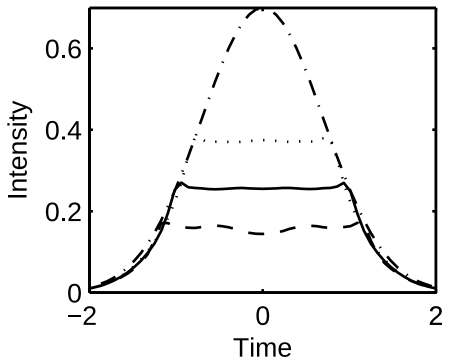

Figure 2.14: Influence of the ratio x of the nonlinearity coefficients of the self-defocusing and self-focusing waveguide on the pulse response of the cascaded asymmetric NLDCs. The dashed line, solid line, and dotted line correspond to x equal to 2, 1 and 0.5 respectively.

$\kappa L_D \geq 1000$, can be satisfied in silica glass if the input pulse is longer than 1 ps. By appropriate doping in glass, it is possible to fabricate the asymmetric NLDC composed of one self-focusing waveguide and one self-defocusing waveguide.

2.4.5 Conclusion

We have demonstrated the limiting property of a mismatched nonlinear directional coupler (NLDC) composed of a self-focusing and a self-defocusing waveguide for both continuous wave and pulse cases and analyzed the working conditions required. The limiting feature can be improved by cascading a few NLDCs together. The limiting threshold and the output power can be adjusted by varying the coupler length or the ratio of the nonlinearity coefficients of the self-defocusing and self-focusing waveguides. In the case of a pulse input, a flat-top pulse is generated when the peak power of the incident pulse exceeds the limiting threshold power, but almost lossless linear behavior is shown as long as the peak power is below the limiting threshold. Unlike in a conventional NLDC, in both cases the rising and falling edges of the output pulse coincide with the input pulse. There is nearly no pulse breakup. Numerical analysis shows that, to achieve a flat top pulse, a coupler coefficient $\kappa L_D > 1000$ is required, and a high IMD could deteriorate the top of the pulse even when the second order dispersion can be ignored. The asymmetric NDLC with $\beta_2 > 0$ shows a better limiting property than the case of $\beta_2 < 0$.

Chapter 3

Optical signal processing by semiconductor optical amplifiers

Within the past decade, optical fiber telecommunication has been the dominant technology in the information transfer domain for its ability to transport large amounts of data, at high data ratio, over a long distance, but at a cheap price [1]. Optical transmission systems with capacities of hundreds of gigabits per second (Gbit/s) are available commercially today, and experimental systems with capacities of several terabits per second (Tbit/s) on a single fiber have been demonstrated in the laboratory [49,50]. Wavelength Division Multiplexing (WDM) is the main technology to achieve such high capacities by carrying many optical channels with different wavelengths simultaneously on the same fiber. Single-channel bit rates have also steadily increased with time, both commercially and in research laboratories. Transmission speeds approaching 0.6 Tbit/s have been demonstrated [51].

However, when data rates are higher than 40 Gbit/s, the transmission systems often involve the use of nonlinear optical devices to access the data, because the direct electronic detection can not respond at so high speed. This is the so-called *electronic bottleneck* [3]. Therefore, lots of network functional devices based on optical nonlinearity should be developed to replace the previous electronic devices, e.g. switching, memory, logic gate, and frequency (wavelength) conversion, etc. In the future, one expects that a completely photonic network will arise as the worldwide data transmission soars up. Semiconductor based devices are widely recognized as prime candidates for photonic networks because of their high nonlinearity, compact size, mature integration technology, low optical and electric power consumption [52].

In this chapter, we will describe the main ideas of the physics of semiconductors: energy bands, electronic wave functions, and optical properties, etc. Next we discuss one specific semiconductor device: the semiconductor optical amplifier (SOA), including its gain and refractive index. The polarization dependence of its gain will also be analyzed. Last we will discuss the mechanisms of the optical signal processing (OSP) in SOA-based devices, and SOA nonlinearities, etc. Most theory of these topics has been extensively described in the literature; however, it is useful to present it here integrally for consistency in the notation, for the use in the following chapters and the future references. Also some discussion of and comments on the existing theory are presented.

Section 3.1 covers the basic ideas behind the $\vec{k}\cdot\vec{p}$ approximation, semiconductor electron band structures and wave functions of bulk semiconductors with and without strain. Section 3.2 gives the optical properties of a semiconductor, focusing on the polarization dependence of optical transitions. Section 3.3 shows the gain and refractive index of SOAs, and analyzes their polarization dependence. Section 3.4 discusses the mechanisms of the optical signal processing by a SOA-based device and the nonlinearities of SOAs.

3.1 Electronic structure of a semiconductor

Fundamental knowledge of the electronic structure of a semiconductor is important for the understanding of semiconductor optical properties and many applications associated with these. Several approaches to describe the band structure of semiconductors exist in the literature [53–57]. We choose the $\vec{k}\cdot\vec{p}$ approximation method [57], which is very convenient to investigate the semiconductor band structure in the vicinity of band edge in the electron $\vec{k}$-space.

3.1.1 The $\vec{k}\cdot\vec{p}$ method

Assuming that one knows the band energies and eigen states at the momentum state $\vec{k}_0$, the $\vec{k}\cdot\vec{p}$ method can be used to explore the band energies and eigen states in the vicinity of $\vec{k}_0$. For direct band gap semiconductors, the optical properties are determined by a small region of states around the Γ point (the position $\vec{k}=0$). Therefore, once the parameters of electron states at $\vec{k}=0$, such as effective masses and energy gaps, are determined experimentally, the $\vec{k}\cdot\vec{p}$ method can efficiently calculate the band energies and eigen states around

$\vec{k} = 0$. This is exactly the interesting region for the investigation of the optical properties of semiconductors.

First we have a look at the one-electron Schrödinger equation. Without electron spin-orbit interaction, the one-electron Schrödinger equation is written as [58]

$$\left[\frac{p^2}{2m_0} + V(\vec{r})\right] \psi(\vec{r}) = E\psi(\vec{r}), \tag{3.1}$$

where m_0 is the free electron mass, $p = -i\hbar\nabla$ is the momentum operator, $V(\vec{r})$ is the crystal potential, $\psi(\vec{r})$ is the electron wave function, E is the energy of the electron, and $\vec{r}$ is the position vector.

The eigenfunctions of Eq.(3.1) can be written in the form of *Bloch waves*, and be specified using two indices: the band index n and the wave vector $\vec{k}$:

$$\psi_{n,\vec{k}}(\vec{r}) = \frac{1}{\sqrt{\Omega}} e^{i\vec{k}\cdot\vec{r}} u_{n,\vec{k}}(\vec{r}), \tag{3.2}$$

where Ω is the volume of the crystal and $\psi(\vec{r})$ is normalized to unity. The Bloch wave, Eq.(3.2), is composed of a slowly varying phase factor, $e^{i\vec{k}\cdot\vec{r}}$, on the scale of a unit cell, and a rapid varying periodic cell function, $u_{n,\vec{k}}(\vec{r})$.

Inserting the Bloch wave, Eq.(3.2), into the one-electron Schrödinger equation Eq.(3.1), we obtain equations for the cell periodic functions $u_{n,\vec{k}}(\vec{r})$

$$\left[\frac{p^2}{2m_0} + \frac{\hbar}{m_0}\vec{k}\cdot\vec{p} + \frac{\hbar^2 k^2}{2m_0} + V(\vec{r})\right] u_{n,\vec{k}}(\vec{r}) = E_n u_{n,\vec{k}}(\vec{r}). \tag{3.3}$$

For each $\vec{k}$, the eigenvalue equation Eq.(3.3) has a complete set of eigen solutions, labelled by n, which completely span the space of periodic cell functions. If the eigenfunction at wave vector $\vec{k} = 0$ is $u_{n,\vec{k}=0}(\vec{r})$, with the eigenvalues $E_n(\vec{k} = 0)$, solutions at other $\vec{k}$ can be expanded in terms of $u_{n,\vec{k}=0}(\vec{r})$:

$$u_{n,\vec{k}}(\vec{r}) = \sum_{n'} C_{n',n} u_{n',\vec{k}=0}(\vec{r}). \tag{3.4}$$

The expansion Eq.(3.4) is inserted into Eq.(3.3) and multiplied by $u^*_{l,\vec{k}=0}(\vec{r})$ from the left. After integration over the whole crystal and using the relation

$$\frac{1}{\Omega_c} \int_{\Omega_c} \mathrm{d}^3\vec{r}\; u^*_{l,\vec{k}=0}(\vec{r}) \cdot u_{n,\vec{k}=0}(\vec{r}) = \delta_{ln}, \tag{3.5}$$

with Ω_c the volume of a unit cell, a matrix eigenvalue equation for the expansion coefficients $C_{n,n'}$ results:

$$\sum_{n'}\left\{\left[E_l(\vec{k}=0)+\frac{\hbar^2k^2}{2m_0}\right]\delta_{l,n'}+\frac{\hbar}{m_0}\vec{k}\cdot\vec{p}_{ln'}\right\}C_{n',n}=E_n(\vec{k})C_{l,n}, \tag{3.6}$$

where

$$\vec{p}_{ln'}=\frac{1}{\Omega_c}\int_{\Omega_c}\mathrm{d}^3\vec{r}\;u^*_{l,\vec{k}=0}(\vec{r})\vec{p}u_{n,\vec{k}=0}(\vec{r}). \tag{3.7}$$

Hence, knowledge of the band edge energy level $E_l(\vec{k}=0)$ and momentum matrix elements $\vec{p}_{ln'}$ is sufficient for the solution of the eigenvalue problem Eq.(3.6), yielding $E_n(\vec{k})$.

3.1.2 Truncation of the space to X, Y, Z, S

The set of functions $u_{n,\vec{k}=0}(\vec{r})$ in Eq.(3.4) span a complete Hilbert space; hence the expansion of $u_{n,\vec{k}}(\vec{r})$ is exact. However, for practical calculations and to obtain a transparent picture, one has to make a selection of a subspace of the most important states. If one is able to select a subspace, with a small number of states, which contains the target state and of which the states have only weak interaction with the states outside that subspace, then an accurate solution can be obtained by diagonalization within the subspace and subsequent inclusion of the interaction with the states outside the subspace by perturbation theory [59].

In a direct band gap semiconductor, the optical property is determined by the states at the band edge with $\vec{k}=0$. The valence band maximum is derived from degenerate atomic p functions, denoted with X, Y and Z states. The conduction band minimum is derived from atomic s functions and is denoted with S state. These four states, X, Y, Z and S form a subspace and are of the main concern in the description of the electron system in semiconductors. The other remote states, derived from higher spherical harmonics, d and f, will be the states outside the subspace and accounted for in a perturbation approach.

The construction of the Hamiltonian H_0, for the $\vec{k}\cdot\vec{p}$ interaction matrix in the basis of S, X, Y and Z states, involves matrix elements

$$\langle\alpha|\,\vec{k}\cdot\vec{p}\,|\beta\rangle\,,\ \text{with}\ \alpha,\beta\in\{X,Y,Z,S\}. \tag{3.8}$$

Considering the symmetry properties of these states: S state is even in all directions, X state is odd in the coordinate x and even in the coordinates y and z, and similarly for Y and Z states, we obtain

$$\langle S|\, p_\alpha \,|\beta\rangle = \delta_{\alpha,\beta}\frac{im_0 P}{\hbar}, \text{ with } \alpha,\beta \in \{X,Y,Z\} \tag{3.9}$$

and

$$\langle \alpha|\, p_\gamma \,|\beta\rangle = 0, \text{ with } \alpha,\beta,\gamma \in \{X,Y,Z\}, \tag{3.10}$$

where P is the only parameter, which should be determined by experimental measurements.

Assuming that the band edge energies of the s and p states are E_s and E_p and using Eq.(3.6), now including the $\vec{k}\cdot\vec{p}$ interaction, the Hamiltonian matrix in the basis of S, X, Y and Z states is obtained as

$$H_0 = \begin{pmatrix} E_s + \frac{\hbar^2 k^2}{2m_0} & iPk_x & iPk_y & iPk_z \\ -iPk_x & E_p + \frac{\hbar^2 k^2}{2m_0} & 0 & 0 \\ -iPk_y & 0 & E_p + \frac{\hbar^2 k^2}{2m_0} & 0 \\ -iPk_z & 0 & 0 & E_p + \frac{\hbar^2 k^2}{2m_0} \end{pmatrix}. \tag{3.11}$$

The secular equation, $|H_0 - E_n(\vec{k})I| = 0$ is solved with I the (4×4) identity matrix and the eigenvalues of Eq.(3.11) are written as

$$\begin{aligned} E_{1,2}(\vec{k}) &= E_p + \frac{\hbar^2 k^2}{2m_0}, \\ E_{3,4}(\vec{k}) &= \frac{E_p + E_s}{2} + \frac{\hbar^2 k^2}{2m_0} \pm \frac{1}{2}\sqrt{(E_s - E_p)^2 + 4P^2k^2}. \end{aligned} \tag{3.12}$$

However, as shown in Fig.(3.1a), the present model is not capable to bend all three valence bands downward; the lack of a band gap is unrealistic. Hence, inclusion of the states outside the S, X, Y and Z subspace is necessary.

3.1.3 Inclusion of the interaction with remote states

The effect of the remote states is included by perturbation theory up to order $\vec{k}^2$, as an effective Hamiltonian within the small subspace, by adding a matrix H_{rem} to H_0, of the form

$$H_{rem} = \frac{\hbar^2}{m_0^2}\sum_{m\neq i,f}\frac{\langle u_{f,\vec{k}=0}|\, \vec{k}\cdot\vec{p}\,|u_{m,\vec{k}=0}\rangle\,\langle u_{m,\vec{k}=0}|\, \vec{k}\cdot\vec{p}\,|u_{i,\vec{k}=0}\rangle}{(E_i + E_f)/2 - E_m} \tag{3.13}$$

where $u_{i,f,\vec{k}=0} \in \{S,X,Y,Z\}$ subspace, and the summation is over all the remote states $u_{m,\vec{k}=0}$. This could be understood as an indirect coupling of states $u_{i,\vec{k}=0}$ to $u_{f,\vec{k}=0}$ via the remote state $u_{m,\vec{k}=0}$.

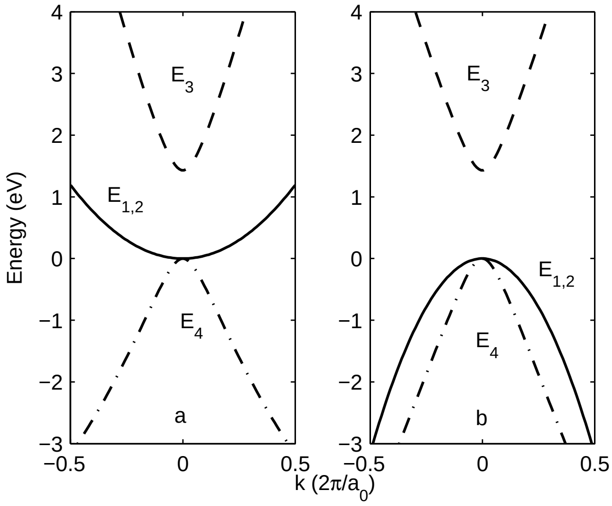

Figure 3.1: (a) Band structure for bulk GaAs, obtained by Eq.(3.12). Note that not all three valence bands bend downward. (b) Band structure for bulk GaAs, obtained by Eq.(3.14), the band gap is now present.

Due to the symmetry properties of these states, a small number of independent parameters expressed by the matrix elements can be used to describe the total $\vec{k} \cdot \vec{p}$ interaction of the S, X, Y and Z states via the remote states.

$$H_{rem} = \begin{pmatrix} Sk^2 & Vk_ykk_z & Vk_zk_x & Vk_xk_y \\ Vk_ykz & Lk_x^2 + M(k_y^2 + k_z^2) & Nk_xk_y & Nk_xk_z \\ Vk_zk_x & Nk_xkk_y & Lk_y^2 + M(k_z^2 + k_x^2) & Nk_yk_z \\ Vk_xk_u & Nk_xk_z & Nk_yk_z & Lk_z^2 + M(k_x^2 + k_y^2) \end{pmatrix}$$

The parameter V is exactly zero in the diamond structure. In the zincblende lattice, V is so small that it is neglected normally[1].

The secular equation for the total Hamiltonian, $H = H_0 + H_{rem}$, is solved for $\vec{k} \in \{k_x, k_y, k_z\}$. Here k_x, k_y and k_z correspondent to the crystal index, and

---

[1] This V parameter contributes to the intrinsic anisotropic properties of semiconductors with zincblende structure.

the band energies are

$$\begin{aligned} E_{1,2}(\vec{k}) &= E_p + \frac{\hbar^2 k^2}{2m_0} + Mk^2 \\ E_{3,4}(\vec{k}) &= \frac{E_p + E_s}{2} + \frac{\hbar^2 k^2}{2m_0} + \frac{S+L}{2}k^2 \pm \\ &\quad \frac{1}{2}\sqrt{\left[E_s - E_p + (S-L)k^2\right]^2 + 4P^2k^2} \end{aligned} \tag{3.14}$$

With a measurement of the zone-center effective masses, the optical matrix element P and Eq.(3.14), a complete description of the band structure can be achieved. We should notice here that V and N are absent in Eq.(3.14), but they will appear for $\vec{k}$ in a general direction. This anisotropic nature of the band structure is a consequence of the interaction with the remote states[2]. However, we will not include this anisotropy in the later chapter.

Until now, a reasonable description of electron structure in the bulk semiconductor is achieved with including the remote states. A comparison of band structures is shown in Fig. 3.1 for cases with and without remote bands. The expected band gap is now present.

3.1.4 Spin-orbit interaction

The electron spin effects, such as spin-orbit interaction, are not yet included. The spin-orbit interaction in vacuum is described to a good approximation by the operator [57]

$$H_{so} = \frac{\hbar}{4m_0^2 c_0^3}(\vec{\nabla} V(\vec{r}) \times \vec{p}) \cdot \vec{\sigma}, \tag{3.15}$$

where $\vec{\sigma}$ are the Pauli spin operators, and c_0 is the speed of light in vacuum. In a semiconductor an expression of similar form arises form the coupling of s and p waves by the $\vec{k} \cdot \vec{p}$ Hamiltonian. This actually dominates over the spin-orbit splitting in vacuum and is accounted for with replacing Eq.(3.15) by an empirical one, e.g. the Luttinger-Kohn Hamiltonian with parameters.

The spin-orbit interaction is treated as a perturbation, by extending the perturbation Hamiltonian to $H_{\vec{k}\cdot\vec{p}} = H_0 + H_{rem} + H_{so}$. We note here, since the coupling of the sp_3 is not included in a obvious way, but simply by parameters, the influence of the sp_3 on the spin is not taken into account in this thesis. It could be included by more complicated 8-by-8 matrix to replace our simple 4-by-4 matrix in section 3.1.5.

---

[2] The remote bands can be included by the parameters from the experimental data.

The Pauli spin operators commute with the non-relativistic Hamiltonian, hence the spin wave function space is just a multiplication of the Bloch functions with the spin eigenfunctions up and down $|\uparrow\rangle$ and $|\downarrow\rangle$. The properties of the Pauli spin operators are summarized as

$$\begin{aligned} \sigma_x|\uparrow\rangle = |\downarrow\rangle, \sigma_y|\uparrow\rangle = i|\downarrow\rangle, \sigma_z|\uparrow\rangle = |\uparrow\rangle \\ \sigma_x|\downarrow\rangle = |\uparrow\rangle, \sigma_y|\downarrow\rangle = -i|\uparrow\rangle, \sigma_z|\downarrow\rangle = -|\downarrow\rangle. \end{aligned} \tag{3.16}$$

Therefore, the original S, X, Y and Z four states are extended to eight stats by including the electron spin freedom.

The part of H_{so} that contains the coupling between the s and p basis states with spin contains the spatial part $\vec{\nabla}V \times \vec{p}$. The x-component of this vector may be written as

$$(\vec{\nabla}V \times \vec{p})_x = -i\hbar\left(\frac{\partial V}{\partial y}\frac{\partial}{\partial z} - \frac{\partial V}{\partial z}\frac{\partial}{\partial y}\right) \tag{3.17}$$

and similar for the other components.

This operator leads to an additional coupling between the basis states, that was still absent with Eqs. (3.9) and (3.10). Considering the symmetry properties of s and p states one finds that the non-zero factors that couple the X, Y and Z states are

$$\langle X|(\nabla V \times \vec{p})_y|Z\rangle = \langle Z|(\nabla V \times \vec{p})_y|X\rangle^*, \tag{3.18}$$

and cyclic permutations of x, y and z. There is only coupling among p states, represented by one additional parameter, the spin-orbit splitting Δ, which is defined as

$$\Delta = -i\frac{3\hbar}{4m_0^2c_0^2}\langle X|(\vec{\nabla}V \times \vec{p})_y|Z\rangle. \tag{3.19}$$

The spin-orbit coupling induces a shift of the $\vec{k} = 0$ energy levels of the valence bands. The corresponding wave functions are chosen as a new basis set. Diagonalization of the combined $\vec{k}\cdot\vec{p}$ and spin-obit Hamiltonian

$$H_{\vec{k}\cdot\vec{p}} = H_0 + H_{rem} + H_{so} \tag{3.20}$$

for $\vec{k} = 0$ produces the new basis and energy eigenvalues describing the system. They are listed in Table (3.1). The J and m_J are the quantum numbers of the total angular momentum $\vec{J} = \vec{L} + \hbar\vec{\sigma}/2$ and its projection J_z along the z axis. The shifted band edges are defined as

$$\begin{aligned} E_c &= E_s \\ E_v &= E_p + \Delta/3 \end{aligned} \tag{3.21}$$

Table 3.1: Eigenvalues and linear combinations of the cell-periodic part of the Bloch function that diagonalizes the total Hamiltonian at $k = 0$.

u_i	$\langle J, m_J\|$		E_i
u_1	$\|\frac{1}{2}, \frac{1}{2}\rangle$	$\|S \uparrow\rangle$	E_c
u_2	$\|\frac{1}{2}, -\frac{1}{2}\rangle$	$\|S \downarrow\rangle$	E_c
u_3	$\|\frac{3}{2}, \frac{3}{2}\rangle$	$\frac{1}{\sqrt{2}}\|(X + iY) \uparrow\rangle$	E_v
u_4	$\|\frac{3}{2}, -\frac{3}{2}\rangle$	$\frac{1}{\sqrt{2}}\|(X - iY) \downarrow\rangle$	E_v
u_5	$\|\frac{3}{2}, \frac{1}{2}\rangle$	$\frac{1}{\sqrt{6}}\|(X - iY) \downarrow\rangle - \sqrt{\frac{2}{3}}\|Z \uparrow\rangle$	E_v
u_6	$\|\frac{3}{2}, -\frac{1}{2}\rangle$	$-\frac{1}{\sqrt{6}}\|(X + iY) \uparrow\rangle - \sqrt{\frac{2}{3}}\|Z \downarrow\rangle$	E_v
u_7	$\|\frac{1}{2}, \frac{1}{2}\rangle$	$\frac{1}{\sqrt{3}}\|(X + iY) \downarrow\rangle + \frac{1}{\sqrt{3}}\|Z \uparrow\rangle$	E_{so}
u_8	$\|\frac{1}{2}, -\frac{1}{2}\rangle$	$-\frac{1}{\sqrt{3}}\|(X - iY) \uparrow\rangle + \frac{1}{\sqrt{3}}\|Z \downarrow\rangle$	E_{so}

and the zero of energy is chosen at E_v, the top of the $J = 3/2$ valence bands. The states u_7 and u_8 are called split-off states because their energy is shifted an amount of Δ by the spin orbit interaction: $E_{so} = E_p + \Delta$.

The total Hamiltonian can be transformed in the $|J, m_J\rangle$ representation by the use of Table 3.1. Diagonalization gives $E(\vec{k})$ for the six valence bands and the two conduction bands. In principle this calculation has be to done numerically. However a simple approximation, with phenomenological parameters Δ, A and B can give quite accurate analytical results, for small k.

3.1.5 Wave functions and energy bands

For the transparency of theoretical derivations it is always an advantage to work with analytical expressions for the electronic energy bands and wave functions. In this section we will describe the semiconductor electron energy bands and wave functions by adopting the Luttinger-Kohn model [60]. First, the material structure parameter V is ignored; thereby the zincblende structure degenerates to the diamond structure and the energy bands become doubly degenerate for each value of $\vec{k}$.

To know the optical properties of the semiconductor materials, only electron states near the band edges are required. Here $\vec{k}$ is so small that the band gap energy $E_g \gg kP$ with $P^2/m_0 \simeq 20$ eV for most semiconductors. This means that the conduction band can be considered isolated from the valence bands and treated as a two-state space. The six valence bands are accounted for via perturbation, and the remote states are neglected, their influence on the band energies being incorporated in the effective parameters of the model. By this approximation, we get

$$E_c = E_c + \frac{\hbar^2 k^2}{2m_0} + \frac{\hbar^2}{m_0^2}\sum_i \frac{|\vec{k}\left\langle u_c\right|\vec{p}\left|u_{vi}\right\rangle|^2}{E_{c,\vec{k}=0} - E_{vi,\vec{k}=0}} + \ldots, \tag{3.22}$$

for the energy, and wave functions

$$u_{cj,\vec{k}} = u_{cj,\vec{k}} + \frac{\hbar}{m_0}\sum_{vi} \frac{\vec{k}\left\langle u_{vi}\right|\vec{p}\left|u_c j\right\rangle}{E_{c,\vec{k}=0} - E_{vi,\vec{k}=0}} + \ldots \tag{3.23}$$

with $j = 1, 2$ and $i = 1..6$ for the periodic part of the wave function.

By definition of an effective mass m_c^* as

$$\frac{1}{m_c^*} = \frac{1}{m_0} + \frac{2P^2}{3\hbar^2}\left(\frac{2}{E_g} + \frac{1}{Eg + \Delta}\right) \tag{3.24}$$

the conduction band energy reduces to

$$E_{c1} = E_{c2} = E_c + \frac{\hbar^2 k^2}{2m_c^*} \tag{3.25}$$

and for the periodic cell wave functions one approximates

$$\begin{aligned} u_{c1,\vec{k}} &= u_{c1} \\ u_{c2,\vec{k}} &= u_{c2}, \end{aligned} \tag{3.26}$$

with u_{c1} and u_{c2} as defined in Tab. 3.1. Hence, the whole wave functions of the conduction band electrons become

$$\begin{aligned} \psi_{c1,\vec{k}} &= \frac{1}{\sqrt{\Omega}} e^{i\vec{k}\cdot\vec{r}} \left|1/2,\ 1/2\right\rangle \\ \psi_{c2,\vec{k}} &= \frac{1}{\sqrt{\Omega}} e^{i\vec{k}\cdot\vec{r}} \left|1/2, -1/2\right\rangle. \end{aligned} \tag{3.27}$$

Our focus is on the optical transitions between the conduction band and the nearest valence bands, which are the four $J = 3/2$ valence bands. An analytical expression for the $J = 3/2$ valence bands and wave functions is desirable. In a direct band gap semiconductor, like GaAs, E_g and Δ both are much larger than kP. In that situation, $J = 3/2$ states are relatively isolated, with a strong coupling between each other but a weak coupling with conduction band and other remote bands. Effectively, the 8-by-8 matrix is then converted into a 4-by-4 matrix, which is known as the Luttinger Kohn matrix [7, 60].

The Luttinger-Kohn Hamiltonian matrix based on four basis states $J = 3/2, m_J = 3/2, -1/2, 1/2, -3/2$, can be expressed as

$$H_{LK} = \begin{pmatrix} H_{hh} & b & c & 0 \\ b^* & H_{lh} & 0 & c \\ c^* & 0 & H_{lh} & -b \\ 0 & c^* & -b^* & H_{hh} \end{pmatrix} \tag{3.28}$$

with these definitions

$$\begin{aligned} b &= \sqrt{3}\frac{\hbar^2}{m_0}\gamma_3(k_x - ik_y)k_z, \\ c &= \sqrt{3}\frac{\hbar^2}{2m_0}[\gamma_2(k_x^2 - k_y^2) - i2\gamma_3 k_x k_y], \\ H_{hh} &= \frac{\hbar^2}{2m_0}\left[\gamma_1 k^2 - \gamma_2(2k_z^2 - k_\perp^2)\right], \\ H_{lh} &= \frac{\hbar^2}{2m_0}\left[\gamma_1 k^2 + \gamma_2(2k_z^2 - k_\perp^2)\right], \end{aligned} \tag{3.29}$$

and $k_\perp = (k_x^2 + k_y^2)^{1/2}$, $\gamma_1, \gamma_2, \gamma_3$ are material parameters. They are in practice empirical parameters deduced from experiment.

The analytical solution of the secular equations $|H_{LK} - E_n(k)I| = 0$, with I the (4×4) identity matrix, is

$$E = \frac{\hbar^2}{2m_0}\left\{\gamma_1 k^2 \pm 2\left[\gamma_2^2(k_x^4 + k_y^4 + k_z^4) + (3\gamma_3^2 - \gamma_2^2)(k_x^2 k_y^2 + k_y^2 k_z^2 + k_z^2 k_x^2)\right]^{1/2}\right\},$$

where ± refers to the light-hole (lh) and heavy-hole (hh) bands respectively. The corresponding two-fold degenerate hh and lh wave functions are in a compact matrix multiplication form

$$\begin{pmatrix} \Psi_h^1 \\ \Psi_h^2 \\ \Psi_l^1 \\ \Psi_l^2 \end{pmatrix} = \frac{1}{\sqrt{\Omega N_i}} e^{i\vec{k}\vec{r}} \begin{pmatrix} -b & R_h & 0 & -c^* \\ -c & 0 & R_h & b^* \\ R_l & b^* & c^* & 0 \\ 0 & c & -b & R_l \end{pmatrix} \times \begin{pmatrix} |3/2\rangle \\ |1/2\rangle \\ |-1/2\rangle \\ |-3/2\rangle \end{pmatrix}, \tag{3.30}$$

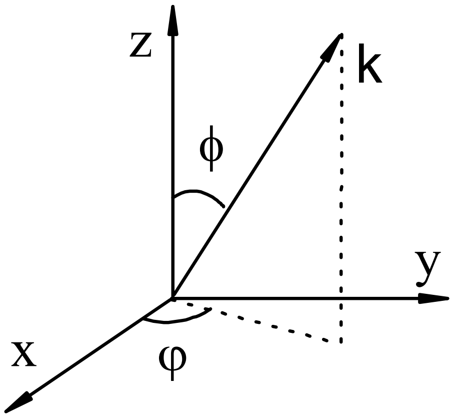

Figure 3.2: Representation of the wave vector for an electron or a hole in the crystalline coordinate system.

with

$$\begin{aligned} R_l &= -H_l + E_l, \\ R_h &= H_h - E_h, \\ N_l &= (R_l^2 + |b|^2 + |c|^2)^{1/2}, \\ N_h &= (R_h^2 + |b|^2 + |c|^2)^{1/2} \end{aligned} \tag{3.31}$$

and E_h and E_l the corresponding band energies.

The wave functions Eq.(3.27) and Eq.(3.30) satisfy the orthogonality condition

$$\langle \psi_{\alpha,n,\vec{k}'} \psi_{\beta,m,\vec{k}} \rangle = \delta_{\alpha,\beta} \delta_{n,m} \delta_{\vec{k},\vec{k}'} \tag{3.32}$$

with $\alpha, \beta \in \{c, lh, hh\}$.

The range of validity of the model is determined by $kP << E_g$ and Δ, as mentioned before. Normally the model is thought to be accurate for electronic wave vectors $k < 0.04(2\pi/a_0)$, which is sufficient in the investigation of the optical properties of semiconductors, since the optical transitions take place close to the band edge.

3.1.6 Uniaxially strained semiconductor

In the previous sections, the electron states in a bulk semiconductor crystal have been investigated. In reality, semiconductor devices are produced by the epitaxial growth technology, thereby the active layer of the device could have

Table 3.2: Material parameters of GaAs

	E_g	Δ	γ_1	γ_2	γ_3	m_{hh}	m_{lh}
GaAs	1.430 eV	0.343 eV	6.85	2.1	2.9	$0.38m_0$	$0.087m_0$

a lattice constant a_e that is different from that of the substrate layer a_s. Such epitaxy is called strained epitaxy and is one of the important emerging areas of crystal growth. The main motivation for such epitaxy is to incorporate strain, which has important effects on the optical properties of the semiconductor. For instance it may reduce the laser threshold current, improve the efficiency and temperature sensitivity, enhance the dynamic response and support high-speed operation [61–65].

For a sufficiently thin active layer, almost all the strain will be incorporated in that layer such that its in-plane lattice constant $a_\parallel$ equals a_s. In the [001] direction, the net strain in the layer plane is expressed as [64]

$$\epsilon_\parallel = \frac{a_s - a_e}{a_e}. \tag{3.33}$$

In response to the biaxial stress, the layer relaxes along the growth direction; the strain $\epsilon_\perp$ being of the opposite sign to $\epsilon_\parallel$. For a tetrahedral semiconductor strained along one of the principal axis, one has $\epsilon_\perp = -\epsilon_\parallel$. The total strain can be represented by a purely axial component $\epsilon_{ax} = \epsilon_\perp - \epsilon_\parallel$ [67].

The Luttinger-Kohn Hamiltonian matrix, also known as Bir-Pikus Hamiltonian [66], for uniaxial strain with the z-axis as symmetry axis, can be expressed in the basis of the four states $J = 3/2, m_J = 3/2, 1/2, -1/2, -3/2$, as [7]

$$H_{LK} = \begin{pmatrix} H_h + \delta & b & c & 0 \\ b^* & H_l - \delta & 0 & c \\ c^* & 0 & H_l - \delta & -b \\ 0 & c^* & -b^* & H_h + \delta \end{pmatrix} \tag{3.34}$$

in which the symbols are defined as in Eq.(3.29). Here, referring to Fig.3.2, $k_z = k\cos\phi$, $k_\perp = k\sin\phi$, with ϕ the polar angle of the electron wave vector, δ is the band splitting energy at $k = 0$ due to the strain. For small uniaxial strain, the band splitting energy varies linearly with the axial strain as $\delta = -b\epsilon_{ax}$, here b is the axial deformation potential [67].

The eigen energies of the matrix are twofold degenerate and given by [64]

$$E = -\frac{\hbar^2}{2m_0}\left\{\gamma_1 k^2 \pm 2\left[(\delta' + \frac{1}{2}\gamma_2(k_z^2 - k_\perp^2))^2 + 3\gamma_3^2 k_z^2 k_\perp^2 + \frac{3}{4}\gamma^2 k_\perp^4\right]^{1/2}\right\} \quad (3.35)$$

where $\delta' = \delta \cdot m_0/\hbar^2$. The energy bands as functions of the electron wave vector, for different angles ϕ of it, are plotted in Fig.3.3. The left of Fig.3.3 is for tensile strain with $\delta = -10$ meV and the right for compressive strain with $\delta = 10$ meV. The material parameters are taken as in Table 3.2.

The curves are for polar angle $\phi = 0, \pi/6, \pi/3$, and $\pi/2$ of the $\vec{k}$ vector; the order is indicated by the arrows. The upper energy will be referred to as E_h and heavy-hole band, the lower as E_l and light-hole band; a nomenclature that does not uniquely correspond to large or small effective mass, as this is obviously an anisotropic tensor in $\vec{k}$ space.

The two eigenvectors corresponding to E_h and those for E_l can be written in a compact matrix multiplication form as superpositions of the m_J basis states with the same form as Eq. (3.30), but with $R_h = H_h + \delta - E_h$, $R_l = -H_l + \delta + E_l$, and $N_i = |R_i|^2 + |c|^2 + |b|^2$, $i = h, l$ and now E_h and E_l the respective eigenenergies Eq.(3.35).

3.2 Optical properties of a semiconductor

Although the optical properties of semiconductor have been extensively investigated in the literature, there are still some interesting effects under discussion, like polarization switching, polarization dependent four wave mixing (FWM), and the photon emission of the electron spin polarized semiconductors. In this section, we recall the principle of the interaction of light and matter, and the fundamentals of the semiconductor optical properties, which will be used in the following chapters.

3.2.1 Interaction of light and matter

In the presence of an external electromagnetic field, the kinetic energy term of an electron is described by a Hamiltonian

$$H = \frac{1}{2m_0}\left[\vec{p} + e\vec{A}(r,t)\right]^2, \quad (3.36)$$

where $\vec{A}(r,t)$ is the vector potential of the electromagnetic field. Choosing the Coulomb gauge and discarding terms with $\vec{A}^2$ one obtains the interaction

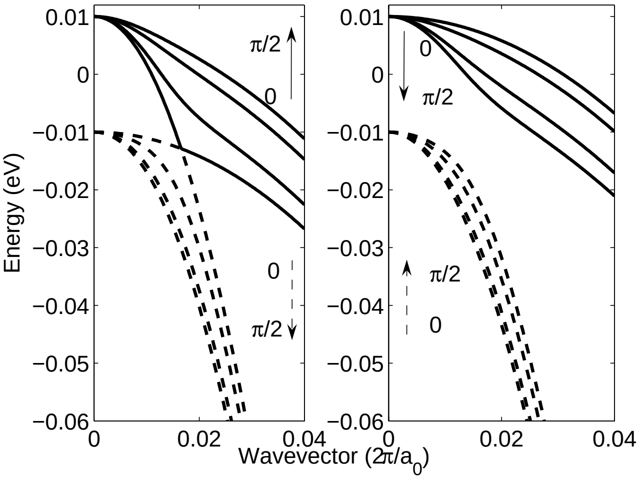

Figure 3.3: Angle ϕ dependent energy band dispersion of E_l and E_h, where ϕ is the angle between the $\vec{k}$-vector of the electron and the strain symmetry axis. The dashed lines are the LH band and the solid lines the HH band. The curves are for $\phi = 0, \pi/6, \pi/3, \pi/2$; the order is denoted by the arrows. (L) tensile strain with $\delta = -10$meV, (R) compressive strain with $\delta = 10$meV.

Hamiltonian of an electron in a radiation field

$$H_{int} = \frac{e}{m_0}\vec{A} \cdot \vec{p}. \tag{3.37}$$

For an electromagnetic field of a given frequency ω the vector potential can be written as

$$\vec{A} = \hat{e}A_0 \exp(i\vec{q} \cdot \vec{r} - \omega t) + c.c., \tag{3.38}$$

where $\hat{e}$ is the unit vector along the direction of the optical field polarization, $\vec{q}$ is the wave vector of the optical field, and we have $\hat{e} \cdot \vec{q} = 0$.

The probability per unit time for a transition induced by a perturbation of the form $He^{\pm i\omega t}$, where H is independent of t, from an initial state $|i\rangle$ of energy E_i to a final state $|f\rangle$ of energy E_f is given by the Fermi Golden rule as

$$\mathrm{P}_{i\rightarrow f} = \frac{2\pi}{\hbar}|\,\langle f|\,H_{int}\,|i\rangle\,|^2\delta(E_f - E_i \pm \hbar\omega). \tag{3.39}$$

Specifically, in a semiconductor the optical transitions occur between conduction bands and valence bands. Therefore, denoting the initial state $|i\rangle$ as valence band state $|\Psi_{v,\vec{k}_i}\rangle$ and the final state $|f\rangle$ as conduction band state $|\Psi_{c,\vec{k}_i}\rangle$, we obtain for the transition probability in semiconductors:

$$\mathrm{P}_{i\rightarrow f} = \frac{2\pi}{\hbar}\left[\frac{eA_0}{2m_0}\right]^2 \delta_{\sigma_i,\sigma_f}|\,\langle\Psi_{c,\vec{k}_f}|\,e^{i\vec{q}\cdot\vec{r}}\hat{e} \cdot \vec{p}|\Psi_{v,\vec{k}_i}\rangle\,|^2\delta(E_f - E_i - \hbar\omega) \tag{3.40}$$

where the spin indices σ_i, σ_f have been included. Eq.(3.40) is the basic expression for the computation of optical constants in the frequency region of interband transitions. The matrix element $\langle\Psi_{c,\vec{k}_f}|\,H\,|\Psi_{v,\vec{k}_i}\rangle$ can be evaluated by insertion of the electron wave functions discussed in section 3.1.5. The usual procedure of integrating over one unit cell and summing over all unit cells results in the condition of conserved momentum in periodic media

$$\vec{k}_f = \vec{k}_i + \vec{q} + \vec{K}, \tag{3.41}$$

where $\vec{K}$ is a reciprocal lattice vector. In the study of semiconductors the typical photon energies are of the order of one eV. The wavelength λ in units of the lattice constant a_0 is then of the order of 10^4. As $\vec{q} = 2\pi/\lambda$ and $\vec{k}_f$ and $\vec{k}_i$ are typically around $2\pi/a_0$, one can safely neglect $\vec{q}$, so adopt

$$\vec{k}_f \simeq \vec{k}_i. \tag{3.42}$$

Hence, only "vertical" transitions in the k-space are allowed. The expression for the probability per unit time is therefore simplified to

$$\mathrm{P}_{i\to f} = \frac{2\pi}{\hbar}\left[\frac{eA_0}{2m_0}\right]^2 \delta_{\sigma_i,\sigma_f} |\hat{e}\cdot\vec{M}_{cv}(\vec{k})|^2 \delta(E_c(\vec{k}) - E_v(\vec{k}) - \hbar\omega), \tag{3.43}$$

where

$$\hat{e}\cdot\vec{M}_{cv}(\vec{k}) = \hat{e}\cdot\int_\Omega \mathrm{d}^3 r \Psi_{c,\vec{k}_f}(\vec{r})\vec{p}\ \Psi_{v,\vec{k}_i}(\vec{r}). \tag{3.44}$$

To obtain the number of transitions $W(\omega)$ per unit time per unit volume induced by light with frequency ω, $\mathrm{P}_{i\to f}$ must be summed over all possible states in the unit volume, i.e. over $\vec{k}$, the spin variable σ, and the band indices v and c. Transforming the summation over $\vec{k}$ into a integral we find

$$W(\omega) = \frac{1}{(2\pi)^3}\frac{2\pi}{\hbar}\left[\frac{eA_0}{2m_0}\right]^2 \sum_{v,c,\sigma}\int \mathrm{d}^3 k \delta_{\sigma_i,\sigma_f} |\hat{e}\cdot\vec{M}_{cv}(\vec{k})|^2 \delta(E_f(\vec{k}) - E_i(\vec{k}) - \hbar\omega). \tag{3.45}$$

The absorption coefficient is by definition the energy absorbed in the unit time in the unit volume divided by the energy flux. Thus we obtain for the absorption coefficient

$$\alpha(\hbar\omega) = \frac{4\pi e^2}{n_r c_0 m_o^2 \omega}\sum_{c,v,\rho}\int \mathrm{d}^3 k \frac{1}{(2\pi)^3} |\hat{e}\cdot\vec{M}_{cv}(\vec{k})|^2 \delta(E_f(\vec{k}) - E_i(\vec{k}) - \hbar\omega). \tag{3.46}$$

The absorption coefficient $\alpha(\hbar\omega)$ can be calculated by Eq.(3.46) with the knowledge of the band structure and the relevant wave functions. Hence, the magnitude, spectral form and polarization dependence of the absorption coefficient is determined by the electronic states of the specific semiconductor structure.

3.2.2 Bulk semiconductor

The optical matrix element $\vec{M}_{cv}$, the conduction band and valence band structures are the prominent ingredients in the calculation of the optical absorption or emission. The optical matrix element is defined as

$$\vec{M}_{cv}(\vec{k}) = \int \mathrm{d}^3 r \Psi_{c,\vec{k}}(\vec{r})\vec{p}\ \Psi_{v,\vec{k}}(\vec{r}). \tag{3.47}$$

The $\vec{k}$ dependence of the Bloch waves has been discussed in section 3.1.1, the conduction band wave function is well described in the single band approximation

$$\Psi_{c,\vec{k}}(\vec{r}) \simeq \frac{1}{\sqrt{\Omega}}\exp(i\vec{k}\cdot\vec{r})u_{c,\vec{k}=0}(\vec{r}). \tag{3.48}$$

Table 3.3: Matrix elements for transitions between the basis states of conduction band and valence band states in units $im_0P/\hbar$

m_J	1/2			-1/2		
Polarization	$\hat{x}$	$\hat{y}$	$\hat{z}$	$\hat{x}$	$\hat{y}$	$\hat{z}$
$3/2$	$1/\sqrt{2}$	$i/\sqrt{2}$	0	0	0	0
$-3/2$	0	0	0	$1/\sqrt{2}$	$-i/\sqrt{2}$	0
$1/2$	0	0	$-\sqrt{2/3}$	$1/\sqrt{6}$	$i/\sqrt{6}$	0
$-1/2$	$-1/\sqrt{6}$	$i/\sqrt{6}$	0	0	0	$-/\sqrt{2/3}$

However, the valence band wave functions are more complicated since they are described in the multi-band formalism as discussed in section 3.1.5. We write the valence band Bloch wave as

$$\Psi_{v,\vec{k}}(\vec{r}) \simeq \frac{1}{\sqrt{\Omega}} \exp(i\vec{k}\cdot\vec{r}) \sum_{vi=1}^{4} f_{vi} u_{vi,\vec{k}=0}(\vec{r}), \tag{3.49}$$

where vi labels the valence band states at $\vec{k}=0$ as defined in Table 3.1, v indicates either light-hole band or heavy-hole band, f_{vi} are the expansion coefficients on the states $u_{vi,\vec{k}=0}$.

Therefore, the optical transition element is written as

$$\begin{aligned} \vec{M}_{cv} &= \frac{1}{\Omega}\int_{\Omega} \mathrm{d}^3 r u^*_{c,\vec{k}=0}(\vec{r})\vec{p}\sum_{vi=1}^{4} f_{vi}u_{vi,\vec{k}=0}(\vec{r}) \\ &= \sum_{vi=1}^{4} f_{vi}\left\langle u_{c,\vec{k}=0}\right|\vec{p}\left|u_{vi,\vec{k}=0}\right\rangle . \end{aligned} \tag{3.50}$$

The matrix elements $\langle u_{c,\vec{k}=0}|\,\vec{p}\,|u_{vi,\vec{k}=0}\rangle$ are calculated from the expressions for the $\vec{k}=0$ Bloch functions in Table 3.1 and are related to the parameter P through the relation $\langle s|\,p_x\,|x\rangle = im_0P/\hbar$ found in section 3.1.2. In summary, we find $\langle u_{c,\vec{k}=0}|\,\vec{p}\,|u_{vi,\vec{k}=0}\rangle = \vec{\xi}\,\langle s|\,p_x\,|x\rangle$, where the components of $\vec{\xi}$ are listed in Table 3.3.

The polarization dependence of transition matrix element $\vec{M}_{cv}$ is determined by the coefficients f_{vi}, which were given by the matrix of Eq.(3.30) for bulk semiconductor. Under non-magnetic conditions, the two HH bands ψ_h^1 and ψ_h^2 are degenerate. One then verifies that

$$\int \mathrm{d}\Omega_k \frac{|R_i|^2}{N_i} = \int \mathrm{d}\Omega_k \frac{|b|^2 + |c|^2}{N_i}, \tag{3.51}$$

where $i = h, l$. This means that the HH band wave functions have equal weights of all four m_J states after integration over all directions of $\vec{k}$, and similarly for the LH band wave functions.

One also easily verifies that

$$\int \mathrm{d}\Omega_k \; cR_h = 0 \text{ and } \int \mathrm{d}\Omega_k \; bR_h = 0 \tag{3.52}$$

and similarly for R_l. This means that all interference terms in the matrix elements $|\vec{M}_{cv}|^2$ disappear upon integration over the angles Ω_k. What remains are just the weight factors f_{vi}. By Eq. (3.51) and (3.52), and Table 3.3, one directly finds the polarization independence of a bulk semiconductor.

This proof, that an isotropic bulk semiconductor medium acts independent of the light polarization, seems here superfluous, as the result follows immediately from symmetry considerations. It is given here, because along the same track one may study the polarization problem in other, special, cases, e.g. strained semiconductors in section 3.2.3, electron spin polarized semiconductors in section 4, or magneto-optics in semiconductors.

In view of the polarization independence of bulk semiconductor and as Eq. (3.51) does not depend on $\vec{k}$, the matrix elements absolute square $|\vec{M}_{cv}(\vec{k})|^2$ can be simply replaced in Eq. (3.46) by $\frac{1}{3}\frac{(m_0 P)^2}{\hbar^2}$, where the factor1/3 stems from the averaging over the three possible linear polarization states $\vec{x}$, $\vec{y}$ and $\vec{z}$ and taking the matrix element out from the integrand of Eq. (3.46). The remaining part is

$$N_{cv}(\hbar\omega) = \frac{2}{(2\pi)^3} \int \mathrm{d}^3 k \delta \left(E_c(\vec{k}) - E_v(\vec{k}) - \hbar\omega \right), \tag{3.53}$$

which defines the joint density of states. It is readily shown that [72]

$$N_{cv}(\hbar\omega) = \sqrt{2} \frac{(m^*)^{3/2} (\hbar\omega - E_g)^{1/2}}{\pi^2 \hbar^3}. \tag{3.54}$$

Therefore, the absorption coefficient of the bulk semiconductor can be written as [72]

$$\alpha(\hbar\omega) = \frac{4\pi e^2}{n_r c_0 m_0^2 \omega} \sum_{v=l,h} \frac{2}{3} \frac{m_0 P}{\hbar^2} \frac{2^{3/2}}{(2\pi)^2} \frac{(m_{cv}^*)^{3/2}}{\hbar^3} \sqrt{\hbar\omega - E_g}, \tag{3.55}$$

where m^*_{cv} is the reduced effective mass: $1/m^*_{cv} = 1/m_c + 1/m_v$; the summation runs over light- and heavy-hole bands.

3.2.3 Uniaxially strained semiconductor

In the previous section, we discussed the absorption in the isotropic bulk semiconductors, where symmetry simplifies the study of their optical properties. This is no longer the case when uniaxial strain is introduced.

In section 3.1.6, we have calculated the band structures and wave functions of semiconductor with uniaxial strain. The optical transition matrix elements squared between the conduction bands and valence bands are $\vec{k}$ dependent and cannot be taken out of the integrand of Eq.(3.46). Also the energy band dispersions $E_c(\vec{k})$ and $E_v(\vec{k})$ are highly anisotropic and the expression for the joint density states N_{cv}, Eq.(3.53), is not applicable any more. The absorption $\alpha(\hbar\omega)$ must then be numerically calculated by Eq. (3.46). This will be further addressed in chapter 4.

3.3 Semiconductor optical amplifier (SOA)

In section 3.1 and 3.2, we examined the electron states and optical properties of a semiconductor. Based on these, we now investigate a specific device, the Semiconductor Optical Amplifier (SOA). A SOA is essentially a semiconductor laser with anti-reflection coated facets that amplifies an injected light signal by means of stimulated emission. Some features of a SOA, like small size, simple electrical pumping, broad spectral range and opportunities for integration and mass production, have made it very promising for signal amplification in a transparent way in all-optical communication systems [68, 69]. In this section, the SOA spontaneous emission spectrum, gain and refractive index will be presented in section 3.3.1. The polarization dependence will be analyzed in 3.3.2.

3.3.1 Gain and refractive index

Lots of applications have been implemented on SOA based devices, in which various linear and nonlinear mechanisms are involved. The essentials of them are just the gain and refractive index of the semiconductors.
The gain of a SOA is often written as [70]

$$g = \Gamma g_m - \alpha_{int}, \tag{3.56}$$

where g is the effective SOA gain coefficient, g_m is the material gain coefficient, α_{int} accounts for various internal losses, Γ is the gain mode confinement factor[3].

The single-pass amplifier gain is

$$G = \exp(gL), \tag{3.57}$$

with L the SOA length.

The material gain g_m can be written in MKS units as [71]

$$g_m(\omega) = \frac{1}{\hbar\omega}\frac{\pi e^2 \hbar}{\epsilon_0 c_0 m_0^2 n_r}\int \mathrm{d}E_{cv}\cdot |M_i|^2 N_{cv}(E_{cv}-E_g)\left(f_c(E_c)-f_v(E_v)\right)\mathcal{L}(E_{cv}) \tag{3.58}$$

where ω is the photon frequency, ϵ_0 the permittivity of free space, c_0 the speed of light in vacuum, m_0 the free electron mass, n_r the refractive index, e the electron charge, $|M_i|^2$ the transition matrix element for linearly polarized light, which can be approximated for bulk semiconductor as [71]

$$|M_i|^2 = |\bar{M}|^2 = \frac{m_0}{6}\left(\frac{m_0}{m_e^*}-1\right)\frac{E_g(E_g+\Delta)}{E_g+2\Delta/3}. \tag{3.59}$$

Here, the influence of the split-off band on the transition matrix is included in the last factor. The joint density of states $N_{cv}(E_{cv}-E_g)$ for bulk semiconductor is Eq. (3.53) with $E_{cv}(k) = E_c(k) - E_v(k) = \hbar\omega$.

The Fermi-Dirac distribution functions for electrons of the conduction and valence bands $f_c(E_c)$ and $f_v(E_v)$ are [72],

$$\begin{aligned} f_c(E_c) &= \frac{1}{\exp[(E_c - E_f^c)/k_B T]+1}, \\ f_v(E_v) &= \frac{1}{\exp[(E_v - E_f^v)/k_B T]+1}. \end{aligned} \tag{3.60}$$

It should be noted that the distribution function for the hole in the valence band is given by $1 - f_v$. The Fermi levels are

$$\begin{aligned} E_f^c &= E_g + k_B T\left[\ln\frac{n_e}{N_c} + \frac{1}{\sqrt{8}}\frac{n_e}{N_c}\right] \\ E_f^v &= -k_B T\left[\ln\frac{n_h}{N_v} + \frac{1}{\sqrt{8}}\frac{n_h}{N_v}\right], \end{aligned} \tag{3.61}$$

---

[3] The confinement factor for gain of guided modes has been shown to be different from the normally used waveguide optical modes confinement factors [78]. However, since we expect that the wave guide mode theory is not sufficient to describe the light propagation in a waveguide shorter than 1mm, we consider the confinement here just as an efficiency factor of the gain

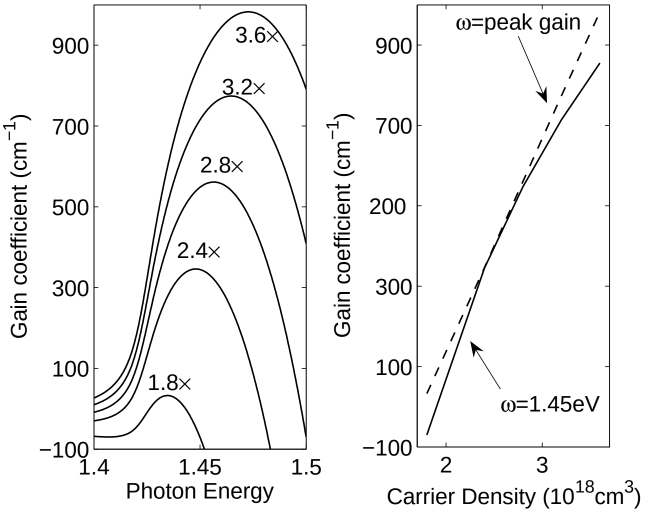

Figure 3.4: (L)Calculated gain spectrum of bulk semiconductor as a function of carrier density. (R) The peak gain has a linear relation with the carrier density, but the relation is nonlinear at a fixed wavelength.

where

$$N_{c,v} = 2\left(\frac{m^*_{e,h} k_B T}{2\pi\hbar^2}\right)^{3/2}. \tag{3.62}$$

Here the valence band edge is adopted as the zero of the energy scale; the conduction band edge energy will be E_g. The emission line broadening is described by the normalized Lorentzian function [73]

$$\mathcal{L}(E_{cv}) = \frac{\hbar/\tau_s}{\pi[(E_{cv} - \hbar\omega)^2 + (\hbar/\tau_s)^2]} \tag{3.63}$$

with τ_s the intraband electron *state* lifetime.

By Eq. (3.58), we calculate the gain spectrum of bulk GaAs in Fig. 3.4. The following parameters are adopted [2]: the GaAs material parameters in Table 3.2, $\tau_s = 100$fs, $\bar{M} = 4.91 m_0$eV.

In semiconductors without rotational symmetry, the integral over the energy, Eq. (3.58), is not convenient any more. Then we take the integral in the $\vec{k}$-space, thus we no longer need the reduced density of states $N_{cv}(E_{cv} - E_g)$ in the left hand side of the equation

$$\int \mathrm{d}E_{cv} \quad \cdot N_{cv}(\hbar\omega - E_g) = 2 \cdot \frac{1}{(2\pi)^3} \int_0^{2\pi} \mathrm{d}\varphi \int_0^{\infty} \mathrm{d}k \int_0^{\pi} \mathrm{d}\phi \cdot k^2 \sin\phi, \tag{3.64}$$

but use the right hand side instead, with the factor of 2 for spin degeneracy. The transition matrix $|M_i|^2$ should also be included in the integrand. By using the uniaxial approximation, we can replace the integral over φ by a factor 2π and the gain is given by

$$\begin{aligned} g_m(\hbar\omega) = 2 \cdot \frac{2\pi}{(2\pi)^3} \frac{\pi e^2 \hbar}{\epsilon_0 c_0 m_0^2 n_r} \frac{1}{\hbar\omega} \int_0^{\infty} \mathrm{d}k \int_0^{\pi} \mathrm{d}\phi \cdot k^2 \sin\phi |M_i|^2 \\ \Big(f_c(E_c(\vec{k})) - f_v(E_v(\vec{k}))) \mathcal{L}(E_{cv} \Big) \end{aligned} \tag{3.65}$$

By Eq. (3.65), we can in principle calculate the gain coefficients of any modified semiconductor structures, like strained bulk semiconductor and quantum well semiconductor.

Similarly the spontaneous emission spectrum can be expressed as [64]

$$R(\omega) = \frac{C_r}{\hbar\omega} \int_0^{\infty} \mathrm{d}k \int_0^{\pi} \mathrm{d}\phi \cdot k^2 \sin\phi |M_i|^2 f_c(1 - f_v) \mathcal{L}(E_{cv}), \tag{3.66}$$

with

$$\begin{aligned} C_r &= 2 \frac{1}{(2\pi)^3} 2\pi \frac{\pi e^2 \hbar}{\epsilon_0 m_0^2 n_r^2} \cdot \rho_o \\ \rho_o &= \frac{1}{\pi^2} \frac{n_r^3}{(\hbar c)^3} (\hbar\omega)^2. \end{aligned}$$

here ρ_o is the optical mode density. $R(\omega)$ is expressed in the MKS units $s^{-1} m^{-3} J^{-1}$.

We must know the relation between the carrier density and the SOA material gain g_m by Eqs. (3.61) and (3.65). Due to the Kramers-Kronig relation, the semiconductor refractive index will also change as the gain changes. In the literature, a parameter, the linewidth enhancement factor α, is defined as

$$\alpha = -\frac{4\pi}{\lambda} \frac{\partial \Delta n_{eff} / \partial N}{\partial g_m / \partial N}, \tag{3.67}$$

with ∂N the carrier density variation, Δn_{eff} the change of the *effective* refractive index of the semiconductors. This definition is very coarse and simple, just a linear relation, and obviously it is based on an assumption that semiconductor the material gain g_m is linearly dependent on the carrier density:

$$g_m \approx A_g(N - N_g), \tag{3.68}$$

with A_g the gain cross section and N_g is the carrier density at transparency. It should be noticed that the linear relation is only valid for the peak gain, not for a fixed frequency, which is usually the situation in optical signal processing. For a fixed frequency, a nonlinear behavior is found, as shown in Fig. 3.4(R).

More precisely the refractive index change should include: carrier band-filling effect through interband transition [74], energy band re-normalization due to free carrier injection [75], and free plasma effects [76]. Moreover, these three effects have a different dependence on the carrier density N. This makes it difficult to formulate a simple model for the the linewidth enhancement factor α.

3.3.2 Anisotropic gain of a SOA [4]

In this section, we discuss a specific and very important problem of the SOA: the polarization dependence of the gain. It is well known that the gain in a SOA is highly anisotropic [68, 77–79]. This anisotropy has been explained intuitively in terms of polarization dependent mode confinement factors caused by the different boundary conditions for the TE and TM fields in the waveguide [68]. But it has been noticed that the measured large gain anisotropy cannot be explained by a weak TE/TM mode confinement factor difference [77, 78]. A confinement factor different from the one based on the Poynting vector, is introduced in [78]: for that case the TE/TM mode confinement factor ratio can be up to 1.4 for an active layer with the minimum thickness[5], and seems to be able to explain the observed large gain anisotropy. We shall demonstrate, however, that although a large confinement factor difference in simulations does yield considerable gain anisotropy, it does not reproduce the significant anisotropy dependence on the pump current that is found experimentally. As we shall show, this dependence is much better reproduced if a weak valence band non-degeneracy is introduced in a heterostructure SOA.

Theoretically the gain anisotropy of a SOA can be expressed in dB as [79]

$$\Delta G = 4.343(\Gamma^{\mathrm{TE}} g_m^{\mathrm{TE}} - \Gamma^{\mathrm{TM}} g_m^{\mathrm{TM}})L \tag{3.69}$$

---

[4] W. Wang, K. Allaart and D. Lenstra, Electron. Lett. **40**. 1602, (2004)

[5] However, this would have some trouble in explaining the case of a quantum well, which has thickness around 10nm, where gain anisotropy is thought to be due to quantum confinement and 1.4 is a too large factor.

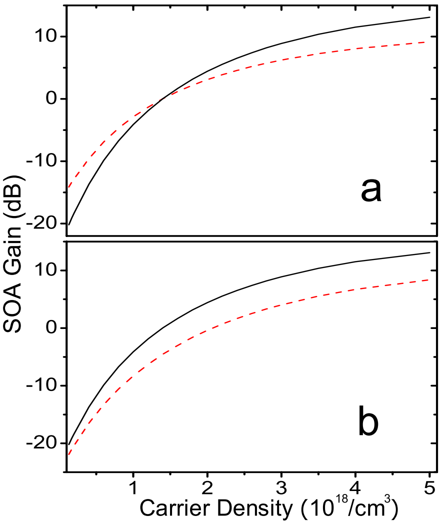

Figure 3.5: Calculated polarization-dependent small-signal gain of TE mode and TM mode as a function of carrier density. The dashed and solid lines are for TM and TE modes respectively. a: polarization anisotropy due to different mode confinement only. b: polarization anisotropy due to intrinsic material gain anisotropy

where Γ^{TE} and Γ^{TM} are the TE/TM mode confinement factors respectively and L is the SOA length. The intrinsic material gains g_m will depend on the carrier concentrations that give rise to emission and absorption of TE and TM polarized light.

From the equations it is obvious that in the absence of intrinsic gain anisotropy the condition $\Gamma^{\mathrm{TE}} > \Gamma^{\mathrm{TM}}$, that is generally applicable, implies positive ΔG above the transparency point. It also implies then negative ΔG below the transparency point. This is illustrated in Fig. 3.5a, where the SOA gain is calculated as a function of the carrier density. The model confinement factor is fitted to the experimentally observed gain anisotropy at large gain in [79]. Note that at the transparency point, $g_m = 0$, the TE and TM gain curves cross: $\Delta G = 0$.

However, to the best of our knowledge, these characteristics are not observed in experiments. The gain anisotropy in [68,77,79] is always positive; no crossing point at transparency is found, while significant saturation of the anisotropy

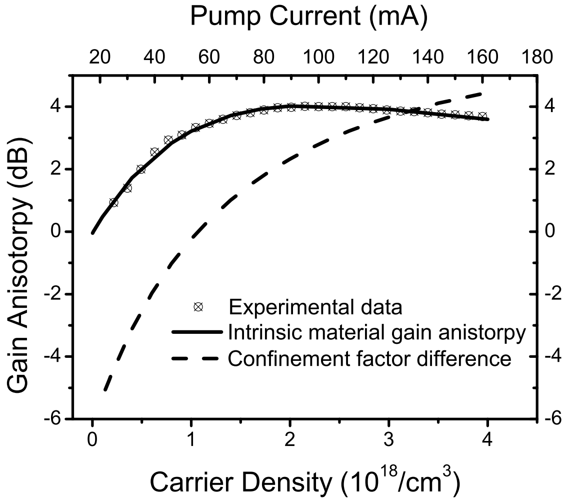

Figure 3.6: Measured gain anisotropy as a function of the pump current, compared with calculated gain anisotropy as a function of the carrier density for a mode confinement factor difference $\Gamma^{\mathrm{TE}} = 0.7\ \Gamma^{\mathrm{TM}}$, and for intrinsic gain anisotropy represented as TE/TM bands splitting of 5 meV.

under high pump current condition is observed.

Since even a bulk SOA is in practice a heterostructure diode, a slight intrinsic anisotropy may still be present in the active layers and give rise to polarization dependence of the gain. In Fig. 3.5b, we calculate the TE/TM polarized light gain in case of 5meV band splitting, corresponding to a practically negligible lattice strain [65]. The results show a similar trend as in [68, 77, 79]. Since the carrier density dependence of the gain anisotropy, caused by the intrinsic material anisotropy is markedly different from that due to different optical model confinement factors Γ^{TE} and Γ^{TM}, we compare the two different simulations with the experimental trend deduced from the data of [79] in Fig. 3.6.

Evidently the model based on the intrinsic anisotropy reproduces the observed trend quite well. In the lower carrier density regime the anisotropy is an almost linear function of the carrier density. As the latter increases, the TE/TM characterized bands become increasingly populated and the gain anisotropy saturates. The slight decrease of the gain anisotropy before saturation occurs when the TE characterized band is totally filled with carriers while the carrier population of the TM characterized band still increases and the band gap shrinks as carrier density increases. If the observed gain anisotropy were due to a difference in mode confinement factors only, the more parabolic trend shown in Fig. 3.6 would have been observed and the gain anisotropy would increase from -6 to 5 dB. In addition, we find that a fit to the experimental data does not allow a confinement factor anisotropy larger than a few per cent.

3.4 Optical signal processing (OSP) based on SOA

In the previous section, we have shown that the semiconductor optical amplifier can be used in optical communication to amplify a weak signal by stimulated emission. However, up to now, the erbium doped fiber amplifier (EDFA), not the SOA, has been the amplifier of choice for optical communication systems, because of the poor performance of gain, noise and nonlinearities of a SOA.

Still, there is much interest in the SOA, owing to its potential functional applications in future all-optical transparent networks. In these applications, the data signal is processed in optical form, rather than first being converted to an electrical signal. As the data rate increases, the electronics bottleneck will require the optical signal data to be processed totally in the optical domain.

In this section, the mechanism of optical signal processing in SOA-based devices and the nonlinearities of SOA are described.

3.4.1 Mechanisms of OSP

The principle of the optical signal processing (OSP) in SOA-based devices is its nonlinearity. The optical nonlinearity of a SOA is very large; more than 10^8 times larger than that of an equivalent length of silica fiber, even for a 100 ps optical pulse [3]. It is caused by carrier density variation when it amplifies the input signals. Associated with this change in carrier density is a concomitant gain and refractive index change. The usual mechanisms of OSP by SOA include: Cross gain modulation (XGM), cross phase modulation (XPM), self-phase modulation (SPM), polarization modulation (PM) and four-wave mixing (FWM) [70].

Cross Gain Modulation: The material gain spectrum of a SOA is homogenously broadened. This means that carrier density changes in the amplifier will affect all of the input signals, so it is possible for a controlling optical pulse at one wavelength λ_p to affect the gain of a target optical signal at another wavelength λ_s. This nonlinear mechanism is called XGM. The most basic XGM scenario would be like: the input data signal at wavelength λ_1, and another target CW signal at wavelength λ_2 are launched together into a SOA. The data signal will modulate the carrier density of the SOA through its own gain. This modulation will be observed by the CW probe, and it is thus imprinted with the inverted data pattern. An optical band-pass filter is needed at the output to suppress the original data signal at wavelength λ_1; therefore we obtain the data signal at the target wavelength λ_2. This function is named: Wavelength Converter, which is used extensively in wavelength division multiplexing (WDM) optical systems.

Cross Phase Modulation: The refractive index of the active region of a SOA is dependent on the carrier density and so on the material gain. This implies that when the gain is cross modulated, the refractive index also changes. The relation between XPM and XGM can be linked to the linewidth enhancement factor α parameter, introduced in Eq. (3.67). XPM can also be used to create wavelength converters and other functional devices. However, because XPM only causes phase changes, the SOA must be placed in an interferometric configuration to convert phase changes of the signals into intensity changes, using constructive or destructive interference.

Self Phase Modulation : SPM is similar to the XPM, but no additional controlling pulse is required. The pulse phase variation is introduced by the pulse itself, which enables automatic switching on/off, by change of the optical pulse intensity. Like in XPM an interferometric configuration is required.

Polarization Modulation : PM is based on XPM or SPM, but without the requirement of an interferometric configuration. PM introduces phase variations between two orthogonal polarization modes (superpositions) of incident linearly or circularly polarized optical pulses. Due to the different phase shifts of the two modes, the polarization of the incident pulse will change, depending on its own or another controlling pulse intensity. This manifests itself as polarization rotation [79–81].

Four Wave Mixing : FWM is a *coherent* nonlinear process that can occur in an SOA, involving two optical fields: a strong pump at angular frequency ω_p and a weaker probe at $\omega_s = \omega_p - \Omega$, with the same polarization. The injected fields cause the amplifier gain to be modulated at the beat frequency Ω. This gain modulation in turn gives rise to a new field at $\omega_p + \Omega$. FWM generation in SOAs can be used in many applications including wavelength converters, multiplexers and demultiplexers.

3.4.2 Ultrafast nonlinearities of a SOA

We have listed some OSP mechanisms of a SOA in the previous section. All of them are in principle based on the refractive nonlinearity of the material, namely

$$\Delta n = n_2|E|^2 \tag{3.70}$$

with Δn the refractive index change due to the existence of the light field, n_2 is the nonlinearity coefficient, and E is the electric field strength of the light.

Before discussing the nonlinear processes in a SOA, which involve carrier intraband dynamics, we first have a look at nonlinearity in silica. There are two types of nonlinearities: Optical Stark effects and multi-wave mixing[6]. Both contribute to the nonlinearity coefficient n_2. Here we only remark that the nonlinearity coefficient is determined by the virtual transition matrix elements. This suggests an approach to increase the nonlinearity of materials by modifying the virtual wave functions.

The concept of virtual states does not refer to the state dissipation; they do not belong to the stationary eigenstates of the system but are allowed within the limitation of Heisenberg's uncertainty principle and the coherence time τ_2. In contrast to those, the optical excitation is resonant in semiconductors; the wave functions are inphase in the coherent time τ_2. The electron configuration is therefore not fully specifical by real observable but also by complex amplitudes. The following description of the nonlinearity of a semiconductor is limited to the time scale larger than τ_2. In chapter 6, we will study the processes during the coherence time scale.

The nonlinearity of SOAs induced by carriers is further complicated by the various carrier intraband processes. When an ultrashort optical pulse is injected into an amplifier, conduction band electrons are depleted because of the stimulated emission processes responsible for the amplification of the input pulse. The reduction in the density of conduction band electrons and valence band holes has two consequences. First, the amplifier gain is reduced and, as a consequence of the induced change in carrier density in the active region, the refractive index of the waveguide changes. Conduction band electron recovery[7] will follow after this depletion. Within a few picoseconds, a quasi-equilibrium (Fermi-Dirac) distribution at the lattice temperature is re-established. From this point on, the state of the amplifier, as far as gain and index are concerned, can be characterized by the total carrier (electron/hole) density. Prior to the establishment of a Fermi-Dirac distribution, the carrier distribution is in a non-equilibrium state governed mainly by Phase Space Burning [82, 85, 105, 106], Spectral Hole Burning [86], and carrier heating and cooling [87]. The temporal

---

[6] It includes Raman and Brillouin effects, i.e., one of the waves will be an optical phonon, the other an acoustic phonon wave.

[7] For simplicity, the analysis of holes in the valence band is not included, but this is similar to that for the electrons of the conduction bands

characteristics of the processes, as well as their dependence on operation parameters, such as wavelength and injection current, can be determined through pump-probe measurements which employ femtosecond optical pulses [88]. A generalized rate equation model that accounts well for such measured gain and index dynamics in bulk and quantum well SOAs can be found in [89,90]. Below the 100-fs time scale, instantaneous coherent processes start to play a marked role, that has nothing to do with real carrier motion any more.

3.5 Summary and Discussion

We have described the general theory of electronic and optical properties of semiconductors in section 3.1 and section 3.2, which will be used in the following chapters. In section 3.3 we presented the formalism for the gain and refractive index of a SOA. These are the basic ingredients for the applications of SOA in an optical network. The issue of polarization dependence of a SOA is investigated in section 3.3.2, where we argue that the confinement factor difference of the TE/TM mode of SOA can not be the most important reason for the polarization dependent gain of a SOA.

The mechanisms of the applications of SOA-based devices in the optical signal processing: XGM, XPM, SPM, PM and FWM are shortly introduced in section 3.4.1. The nonlinear processes of a SOA are discussed in section 3.4.2, where we distinguish between coherent and non-coherent processes. For the coherent ones, the wave functions of the electron are in-phased in the coherent time τ_2. The electron configuration is specified by complex amplitudes and virtual populations. For non-coherent processes, the wave functions are de-phased, in some way. One can say that the electrons population becomes real now. Then one can start to study the electron intraband dynamics. The non-coherent processes, described in a phenomenological way, include several aspects of electron intraband dynamics: momentum space burning, spectral hole burning, carrier heating and cooling, and carrier density distribution recovery.

Chapter 4

Optical polarization in electron-spin polarized semiconductors

4.1 Introduction [1]

An electron, apart from charge and mass, has an intrinsic angular momentum, *spin* [7,91]. The interest in semiconductor electron spin comes from the fact that electron spin angular momentum couples to its orbital angular momentum. As we have shown in section 3.1.4, this coupling results in a valence band structure including heavy-hole, light-hole and split-off bands [7]. However, since III-V semiconductors, like GaAs, are normally non-magnetic materials, we shall treat the bands as spin degenerate, though the spin degeneracy is actually removed by, for instance, the Dresselhaus effect. Macroscopically, the electron spin does not show up in the properties of electronic equipment we are using everyday.

Theoretically, it is possible to use spin, as a degree of freedom of the electron in electronic devices for information processing. The proposed devices are called *Spintronics* and they are believed to have the potential to outperform conventional electronic circuits in speed, integration density and power consumption [91]. During recent years, spintronics has been explosively developed in semiconductor physics: spin polarized electron generation, injection, and transport [91–94]. However, since the spin polarized electrons are not in an equilibrium state without a magnetic filed, the spin orientation of electrons

---

[1] W. Wang, K. Allaart and D. Lenstra, Phys. Rev. B **74**. 073201, (2006); selected by Virtual J. of Nanoscale Sci. & Tech. **14**. 8, (2006)

is rather fragile. Therefore, before a form of semiconductor spintronics can be globally applied, it is necessary to understand electron spin dynamics first.

The most efficient way to study the electron spin dynamics is the pump-probe or pump-luminescence technique based on *Optical Orientation* [95]. The principle can be explained by the optical transition selection rules. If circularly polarized light excites a semiconductor, the photo-excited electrons and holes will be spin polarized, with opposite sign. A certain fraction of them will recombine before the spin polarization is lost by spin relaxation processes. By measuring the circular polarization of the luminescence it is therefore possible to study the spin dynamics of the non-equilibrium carriers in semiconductors and to extract such useful quantities as the spin orientation, the recombination time, or the spin relaxation time of the carriers. However all of those techniques presuppose that one knows the exact relation between the luminescence circular polarization P_{cir} and the electron spin polarization P_s.

In this chapter we will rigorously prove the relation $P_{cir} = -\frac{1}{2}P_s \cos\theta$ in case of a bulk semiconductor, with θ the angle between the observation direction and the spin polarization direction. This relation is extensively used in the literature, but the way it is understood is not consistent with the experimental setups [92, 93, 96, 98–100]. Also, we investigate the relation in the uniaxial strained bulk semiconductor. We noticed that, in uniaxially strained bulk semiconductors, it is possible to achieve high spin polarization of the electrons in certain parts of the energy spectrum by circular polarized light. We also suggest a method for detection of strain in bulk semiconductor by measuring its circular polarization of the luminescence. This method is comparable to the method suggested in [101], but should be more precise in the non-magnetic semiconductor.

4.2 Model

The spin polarization P_s of the conduction band electrons is defined as [91, 92, 99, 102]

$$P_s = \frac{N_+ - N_-}{N_+ + N_-}, \tag{4.1}$$

with $N_\pm$ the density of electrons with spin parallel and antiparallel to the magnetization direction. Here the spin states of electron are defined as $|\uparrow\rangle$ and $|\downarrow\rangle$ with respect to a certain macroscopic direction, see Ref. [98]. The circular polarization P_{cir} of the light is defined as

$$P_{cir} = \frac{I_+ - I_-}{I_+ + I_-}, \tag{4.2}$$

with $I_\pm$ the intensities of right(helicity σ^+) and left(σ^-) circular polarized components of the emitted light, defined with respect to the detection direction. The emitted light intensities for photon energy $\hbar\omega$ are determined by the total spontaneous emission rate as we have shown in Eq.(3.66) [64, 72], but here we focus on the *circular* polarization light transition matrix element squared $|M^D_\pm(\vec{k})|^2$, which is defined according to the observation direction, of the detector D. The emission rate is

$$I_\pm(\omega) = C_r \int_0^\infty \mathrm{d}k \int_0^\pi \mathrm{d}\phi \big[k^2 \sin\phi \cdot f_e(E_{\vec{k}}) f_h(E_{\vec{k}}) \\ \cdot \mathcal{L}(E^{cv}_{\vec{k}}) \cdot |M^D_\pm(\vec{k})|^2\big] \tag{4.3}$$

with

$$C_r = \frac{1}{(2\pi)^3} \frac{2e^2 n_r \hbar\omega}{m_0^2 c^3 \hbar^2}. \tag{4.4}$$

The integration is over the wave vector $\vec{k}$ of the charge carriers involved in the transition; ϕ is the polar angle of $\vec{k}$ with respect to the axis of the uniaxial strain, which is also adopted as the quantization axis of the electron spin. $\mathcal{L}(E^{cv}_{\vec{k}})$ is the Lorentzian function as defined in Eq.(3.63) [73], in which $E^{cv}_{\vec{k}}$ is the energy difference between the initial and final band states with wave vector $\vec{k}$. The fermi distributions f_e and f_h for electrons and holes respectively are those for intraband thermal equilibrium, which is reached very fast as compared to the spontaneous emission life times. The label D on the dipole transition matrix element $M^D_\pm(\vec{k})$ indicates that right or left handedness must be defined with respect to the detector direction. All factors in the integrand of Eq.(4.3) are functions of both magnitude of $\vec{k}$ and angle ϕ.

To derive a relation between the electron spin polarization P_s and circular polarization P_{cir}, one must know the energy band structure of a semiconductor. That gives the injected electron distributions $f_e(E_{\vec{k}})$, $f_h(E_{\vec{k}})$, the optical transition energy $E^{cv}_{\vec{k}}$, the line shape function $\mathcal{L}(E^{cv}_{\vec{k}})$, and also the band wave functions that determine $|M^D_\pm(\vec{k})|^2$, and therewith the optical transition probabilities. Therefore, once the band structures and wave functions are known, the relations between the P_{cir} and P_s are determined by Eq.(4.3).

In several theoretical studies of the optical transitions between conduction and valence band, one denotes the heavy-hole (HH) band by $|m_J = \pm 3/2\rangle$ and light-hole (LH) band by $|m_J = \pm 1/2\rangle$, referring to the $J = 3/2$ multiplet at $\vec{k} = 0$ [92, 93, 96, 98–100]. One should be aware that this notation only applies if the magnetic quantum number m_J refers to the projection of the angular momentum J on the direction of the wave vector $\vec{k}$. For that reason the notation m'_J should be preferred, as distinct from m_J for the projection on a fixed laboratory axis, independent of the direction of the wave vector $\vec{k}$. In this work the fixed laboratory z-axis will be the axis of the uniaxial strain, which is also adopted as the quantization axis for the spin polarization of the conduction electrons. From now on m_J refers to the projection of J on that axis.

For a semiconductor in which the rotational symmetry is not broken by strain, the energy of the valence band states depends on the modulus of $\vec{k}$ only, not on its direction $\hat{k}$. Consequently all m_J states are equally represented at a given band energy. Then one may ignore the distinction between m'_J and m_J in the calculation of the optical polarization [103]. However, when the rotational symmetry is broken by strain, the structure of the valence band states and thereby also the strength of the optical transitions, becomes dependent on both the magnitude and the direction of the $\vec{k}$ vector. While in a symmetric structure the relation $P_s = -P_{cir}/2$ holds [91,95,102,103] for HH as well as for LH bands, being an average over the $m_J = \pm 3/2$ as well as over the $m_J = \pm 1/2$ states, we show that even in case of modest strain one may obtain for each band a value of P_{cir} close to P_s or to $-P_s$, depending on whether one has tensile or compressive strain. This results can be compared with the conclusions for a quantum well semiconductor; those that have been reported seem not consistent, however: $P_s = P_{cir}$ in Ref [96], but also, even below 10 K, $P_s \leq 0.6 P_{cir}$ in Ref [97].

Here we point out that, since in case of strain the composition of HH and LH states depends, for a given energy, on both magnitude and direction of $\vec{k}$, the polarization of the luminescence light at a certain photon energy will be photon energy dependent, even after integration over all directions of the $\vec{k}$-vector involved. This variation of the photon polarization with photon energy will be shown to be characteristic for compressive or tensile strain and quite sensitive to strain. It implies also that for the reverse process, the creation of spin polarized conduction electrons by absorption of circularly polarized light, an optimal photon energy range can be identified.

The optical transition probability, Eq. (4.3), is determined not only by the dipole transition matrix element; also the band energy dispersion and dispersion with the angle of the wave vector $\vec{k}$ must be taken into account. This requires a numerical calculation which we perform for an illustrative case with realistic material parameters of GaAs as the active layer of the heterostructure.

The conduction band wave functions are presented in Eq. (3.27), the valence band wave function in Eq. (3.30). Therefore he optical transition matrix

Table 4.1: Optical transition matrix elements, in units $\mu/2\sqrt{3}$, between the conduction band $|\Psi_c\rangle$ and the valence band basis states $|m_J\rangle$ for circularly polarized light detected under angle θ with the electron spin polarization axis.

States	$\langle S\uparrow\| r_-$	$\langle S\uparrow\| r_+$	$\langle S\downarrow\| r_+$	$\langle S\downarrow\| \vec{r}_-$
$\|+3/2\rangle$	$-\sqrt{3}(1+\cos\theta)$	$\sqrt{3}(1-\cos\theta)$	0	0
$\|+1/2\rangle$	$2i\sin\theta$	$2i\sin\theta$	$1-\cos\theta$	$-(1+\cos\theta)$
$\|-1/2\rangle$	$1-\cos\theta$	$-(1+\cos\theta)$	$2i\sin\theta$	$2i\sin\theta$
$\|-3/2\rangle$	0	0	$-\sqrt{3}(1+\cos\theta)$	$\sqrt{3}(1-\cos\theta)$

element between the conduction band and valence band is

$$\langle\Psi_c|\, r_\pm\, |\Psi_h^i\rangle = \sum_{m_J} f^i_{m_J}\cdot\langle\Psi_c|\, r_\pm\, |m_J\rangle\,, \tag{4.5}$$

where the expansion coefficients $f^i_{m_J}$ of the valence band wave function Ψ_h^i on the basis states are given by a row of the matrix in Eq.(3.30); $r_\pm$ is the dipole transition operator for right or left circular polarized light.

We consider $\langle S|\, r_\pm\, |m_J\rangle$ first. If the detection direction $\hat{D}$ makes an angle θ with the z-axis, the detected circularly polarized light σ^+ and σ^- correspond to the optical transition operators $r_\pm$

$$\begin{aligned} r_+ &= -\frac{1}{\sqrt{2}}(x+iy\cos\theta-iz\sin\theta) \\ r_- &= \frac{1}{\sqrt{2}}(x-iy\cos\theta+iz\sin\theta). \end{aligned} \tag{4.6}$$

The transition matrix elements $\langle\Psi_c|\, r_\pm\, |m_J\rangle$ between the basis states are listed in Table I in units: $\mu/2\sqrt{3}$, with $\mu = |\,\langle S|\, x\, |X\rangle\,| = |\,\langle S|\, y\, |Y\rangle\,| = |\,\langle S|\, z\, |Z\rangle\,|$.

4.3 Correlation between P_{cir} and P_s in Bulk Semiconductor

The calculation of the emission rates Eq.(4.3) is complicated by the fact that all factors in the integrand depend on the direction of the $\vec{k}$-vector. A great simplification occurs when the strain splitting δ is equal to zero, i.e. when the material is isotropic. In that case the Fermi function f and the Lorentzian factor

L are independent of the angle ϕ of $\vec{k}$. Then the polarization P_{cir}, Eq. (4.2), is simply determined by the ratio of the squared transition matrix elements $|M^D_{\pm}|^2$ integrated over all directions of $\vec{k}$.

Moreover, as we have shown in section 3.2.2, for the isotropic semiconductor, the integrated weight factors of the different magnetic substates $|m_J\rangle$ are all equal, and the cross terms disappear. Then one can conclude immediately that the circularly polarized light intensities are proportional to the sum of these squared dipole matrix elements in Table 4.1;

$$\begin{aligned} I_{\pm}(E_k) = n_+(E_k)&[3(1 \mp \cos\theta)^2 + 4\sin^2\theta + (1 \pm \cos\theta)^2] \\ +n_-(E_k)&[(1 \mp \cos\theta)^2 + 4\sin^2\theta + 3(1 \pm \cos\theta)^2], \end{aligned} \tag{4.7}$$

where $n_{\pm}(E_k)$ are spin polarized electron densities at energy E_k. The circular polarization P_{cir}, Eq. (4.2), is therefore [102–104]

$$P_{cir}(\omega) = \frac{n_+(E_k) - n_-(E_k)}{n_+(E_k) + n_-(E_k)} \cdot \frac{-\cos\theta}{2} = \frac{-\cos\theta}{2} P_s(E_k). \tag{4.8}$$

This conclusion holds strictly for transitions between the conduction bands $|\Psi_c\rangle$ and valence bands $|m_J\rangle$ in an isotropic semiconductor. When the spin polarization P_s of the electrons is energy independent, also the circular polarization P_{cir} of the emitted light will therefore be wavelength independent. But this is not the case when $\delta \neq 0$. In that case Eq.(3.51) does not apply anymore; the optical transition matrix elements squared between the conduction bands and valence bands cannot be replaced by equal weights of the contributions of the individual basis states as in the symmetric structure. So the exact $f^i(k,\theta)$ must be taken into account in Eq.(4.5). Also the energy band dispersions of E_l and E_h now play a role via the Fermi distribution $f_i(E_k)$ and the gain spectrum $\mathcal{L}(E^{cv}_k)$. The emitted light intensity must therefore in that case be numerically calculated by Eq.(4.3).

4.4 Correlation in Strained bulk semiconductor

As the strain influences all factors in the integrand of Eq.(4.3), we try to elucidate the final result by exposing the effect of the strain anisotropy on some of the factors that play a role in the emission rates separately. We first consider only the transition matrix $|M^D_{\pm}(k)|^2$ of Eq.(4.3), namely the discrimination of emitted light σ^- and σ^+ solely determined by this transition matrix. In Fig. 4.1 we show the magnitude of the components, $|f^h_{m_J}|^2$ and $|f^l_{m_J}|^2$, of the HH and

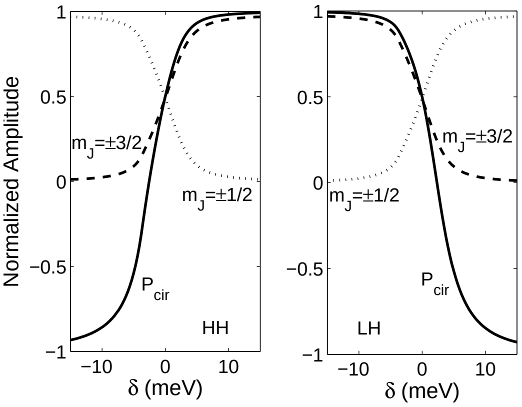

Figure 4.1: Dependence of the components $|f^i_{m_J}|^2$ on the strain δ, for the HH band (L) and LH band (R). The solid line gives the circular polarization of the emitted light $-P_{cir}$ for each of the bands when one considers the factor $|M^D_\pm(k)|^2$ only.

LH bands as a function of band splitting energy δ due to a strain, integrated over all directions of $\vec{k}$, at $k = 0.01a_0$, with $a_0 = 2\pi/a$; a the lattice constant.

In the isotropic case, $\delta = 0$, we have $|f^i|^2 = 1/2$ for $m_J = \pm1/2, \pm3/2$ in both HH and LH bands. When δ increases, i.e. compressive strain, $m_J = \pm3/2$ components of HH band wavefunction and $m_J = \pm1/2$ components of LH band wavefunction increase and saturate. At the same time $m_J = \pm1/2$ components of HH band wavefunction and $m_J = \pm3/2$ components of LH band wavefunction decrease and saturate. Therefore, the circular polarization P_{cir} of the light emitted in the transition to the HH bands will approach to $-P_s$, and for the LH band P_{cir} approaches P_s. When δ is negative, i.e. tensile strain, these results are just opposite to those for compressive strain.

Next, we take the energy band dispersions Eq.(3.35) of E_l and E_h into account. As shown already in Fig. 3.3 the energy band dispersion is significantly dependent on the angle ϕ of the electron momentum vector $\vec{k}$. Consequently also the Fermi distribution, optical transition energy $E^{cv}_{\vec{k}}$ and Lorentzian broadening of the gain spectrum $\mathcal{L}(E^{cv}_{\vec{k}})$ are ϕ dependent. We checked whether it is a good approximation to average this angle ϕ dependence, for the HH and for the LH energy bands, in order to simplify the calculations of the integrals Eq. (4.3). For this purpose, we show in Fig. 4.2 the angle ϕ dependence of the integrand in Eq.(4.3) for $k = 0.02a_0$, $\delta = 5$ meV and for the average

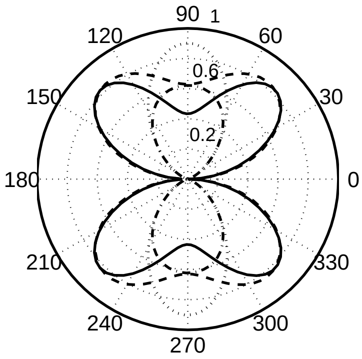

Figure 4.2: Dependence of transition strength son the angle ϕ of the electron momentum $\vec{k}$. The solid and dash-dotted lines are exact calculation for transitions between the conduction bands and $m_J = \pm 3/2$ and $m_J = \pm 1/2$ components of the valence band respectively. The dashed and dotted lines are the corresponding quantities calculated with an angle-averaged energy $\bar{E}_k$.

resonant frequency of the optical transitions. The solid and dash-dotted lines are the exact expressions $f_e(E_{\vec{k}})f_h(E_{\vec{k}})\mathcal{L}(E^{cv}_{\vec{k}})|M^D_\pm(\vec{k})|^2 \sin\theta$ as in Eq.(4.3) for transitions between conduction bands and $|\pm 3/2\rangle$ and $|\pm 1/2\rangle$ components of the HH bands; the dashed and dotted lines are the corresponding quantities calculated with the angle-averaged energy $\bar{E}_k$. We find significant differences even for this small k and weak strain. Therefore the energy band structure can not be simply approximated by one averaged pair of HH and LH bands; the luminescence spectra must be calculated by including the full ϕ dependence of all factors in the integrand of Eq.(4.3).

As an illustrative example we present here the calculated polarization of the luminescence spectrum with realistic parameters for GaAs [72]. The luminescence spectrum is shown in Fig. 4.3 as the dash-dotted line. The solid line is the ratio of the luminescence circular polarization P_{cir} and electron spin polarization P_s, assuming a P_s independent electron energy. We distinguish contributions of $|S\uparrow\rangle$-LH transitions and $|S\uparrow\rangle$-HH transitions to the luminescence spectrum by dotted lines and dashed lines. The upper dashed line and lower dotted lines are transitions stemming from $|S\uparrow\rangle \rightarrow |\pm 3/2\rangle$; the lower dashed line and upper dotted lines from $|S\uparrow\rangle \rightarrow |\pm 1/2\rangle$. The adopted material parameters are from [72], electron density is $10^{18}/\text{cm}^3$, $\tau_s = 100$ fs and $\delta = 10$ meV. The luminescence P_{cir} is found to be photon energy dependent, and in

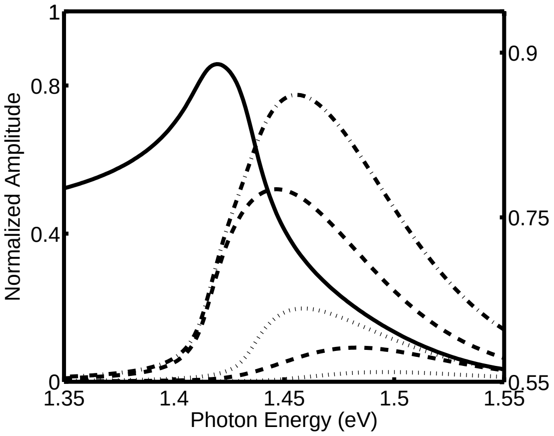

Figure 4.3: Normalized luminescence spectrum and its photon circular polarization in a compressively strained semiconductor with strain energy splitting $\delta = 10$ meV. The dash-dotted line is the luminescence spectrum. The solid line is $-P_{cir}/P_s$ with its scale on the right vertical axis. The dashed and dotted lines are luminescence of $|S\uparrow\rangle$-HH and $|S\uparrow\rangle$-LH transition components respectively. The upper dashed line and lower dotted lines are transitions with optical property of $|S\uparrow\rangle \rightarrow |\pm 3/2\rangle$; the lower dashed line and upper dotted lines with optical property of $|S\uparrow\rangle \rightarrow |\pm 1/2\rangle$.

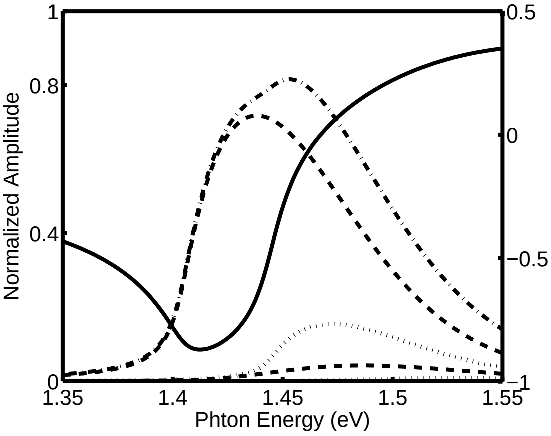

Figure 4.4: Same as Fig. 4, but for a tensile-strained semiconductor with $\delta = -20$ meV. Here upper dashed line and lower dotted lines are transitions with optical property of $|S\uparrow\rangle \rightarrow |\pm 1/2\rangle$; the lower dashed line and upper dotted lines with optical property of $|S\uparrow\rangle \rightarrow |\pm 3/2\rangle$.

the range of $1.41 < \hbar\omega < 1.43$ (eV) high circular polarization P_{cir}, to nearly $-0.9P_s$, can be reached. There is no broad-spectrum rule to govern the relation between P_s and P_{cir} like the simple spectrum independent relation in the isotropic case.

Figure 4.4 shows a luminescence spectrum like Fig. 4.3, but now for tensile strain. One might expect from Fig. 4.2 that the circular polarization P_{cir} will approach P_s . However, even for a large strain energy splitting of $\delta = -20$ meV, the peak value of P_{cir} does not exceed $0.8P_s$; For smaller strain, $\delta = -10$ meV, the maximum of P_{cir} is around $0.5P_s$. This is mainly due to the transition strength for $m_J = \pm 3/2$, which is three times that for a $m_J = \pm 1/2$ state, and also to the smaller effective mass of the LH bands.

One may notice that the strain of the semiconductor has a dramatic influence on the energy dependence of the luminescence circular polarization $P_{cir}(\omega)$. On the other hand, by spectroscopic analysis of $P_{cir}(\omega)$, one is able to know to what extent the material is strained and what is the nature of the strain[2].

In Fig. 4.5 and 4.6 we show polarization correlations of P_s and P_{cir} as a function of the strain with realistic parameters for GaAs in tensile and compressive strained cases, respectively. For tensile strain, with band splitting energy

---

[2] W. Wang, K. Allaart and D. Lenstra; *Proceeding of the 17th Quantum Electronics and Photonics Conference*, Manchester, U.K. Sep. 4-7, 2006

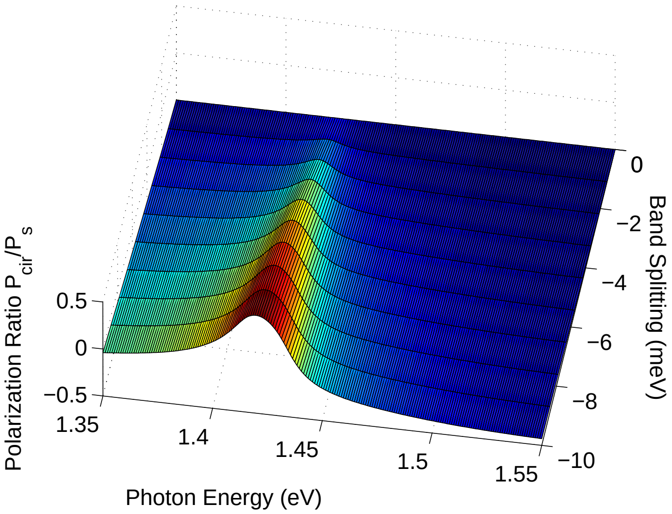

Figure 4.5: The strain dependent correlation between the electron polarization P_s and the circular polarization of the emitted light P_{cir}, in case of tensile strain.

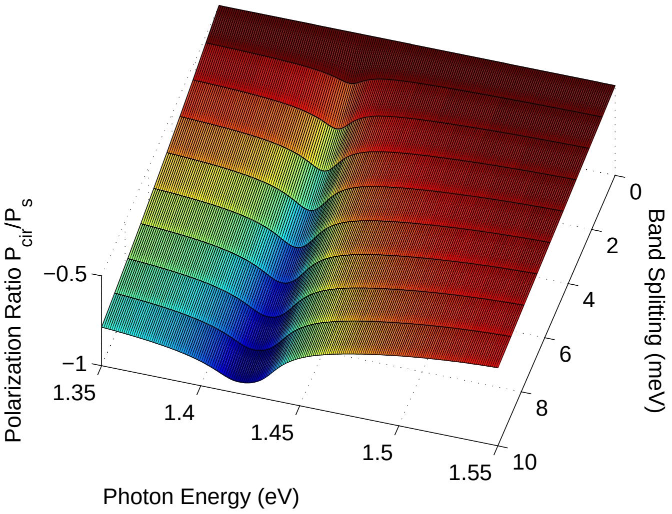

Figure 4.6: The strain dependent correlation between the electron polarization P_s and the circular polarization of the emitted light P_{cir}, in case of compressive strain.

decreasing from 0 to -10 meV, the ratio of P_s and P_{cir} increases from -0.5 to 0.5. The ratio approaches 0.8 at $\delta = -20$ meV. For compressive strain, the ratio between P_s and P_{cir} decreases from -0.5 to -0.9, when the strain induced band splitting energy increases from 0 to 10 meV. Further increase of it makes the ratio approach -1. Comparing Fig. 4.5 and 4.6, one finds significantly different polarization correlation in various strained cases, both regarding the values and the trends as a function of photon energy. This provides a way to investigate the strain in semiconductors by measuring the correlation between the injected electron spin polarization and the luminescence circular polarization.

4.5 Conclusion

We analyzed the correlation between the spin polarization P_s of the injected electrons and the circular polarization P_{cir} of the emitted light in a heterostructure semiconductor with and without internal strain.

When the strain δ is zero, the relation $P_{cir}(\omega) = -\cos\theta/2 \cdot P_s$ holds independent of photon energy $\hbar\omega$ and of the degeneracy of the valence bands. This is essentially based on the property that the twofold degenerate HH as well as LH band wave functions in this case have expansion coefficients of equal weight for the four $|m_J\rangle$ basis states, independent of the energy. When strain $\delta \neq 0$, the luminescence $P_{cir}(\omega)$ is significantly photon energy dependent. In case of $\delta > 0$, i.e. compressive strain, $P_{cir} \simeq -0.9P_s$ can be reached over 100nm width of the spectrum even under weak strain. In case of $\delta < 0$, i.e. tensile strain, P_{cir} does not exceed $0.8P_s$, even under large strain. This is mainly due to the smaller effective mass of LH bands and larger transition matrix elements for $m_J = \pm 3/2$ than for $m_J = \pm 1/2$.

The characteristic photon energy dependence, caused by strain, of the correlation between electron spin polarization P_s and circular photon polarization P_{cir}, provides a spectroscopic tool to investigate the nature of strain in semiconductor material. On the other hand, in the strained semiconductors, the strong correlation between P_{cir} and P_s in a certain range of photon energy enables efficient creation of electron spin polarization by circularly polarized light excitation.

Chapter 5

Mechanism of induced birefringence in semiconductors

In view of promising applications in THz all-optical signal processing, the ultrafast dynamics of SOA gain and refractive index has been studied extensively. Sub-picosecond nonlinearities of gain and refractive index have been observed in experiments and were attributed to the intra-band non-equilibrium dynamics in the carrier energy distribution [86, 87, 105–111]. The main processes are identified to include two photon absorption(TPA), spectral hole burning (SHB) and carrier heating (CH) [88–90].

In some pump-probe experiments an instantaneous and negative phase shift is observed. The explanations in the literature of this phenomenon include: Stark effects, Kerr effects, rapid electronic or some virtual processes, etc [89,108, 111,112]. These explanations focus on the instantaneous change of the refractive index. Some explanations are based on the carrier density intraband dynamics, assuming an ultrafast carrier scattering time and obtain some agreement with the observed phase shift of the probe [81,108]. In the modelling, this phase shift has been attributed to the TPA process with a negative linewidth enhancement factor. From the above explanations the TPA process should not influence the polarization of a probe light polarized $\pi/4$ to an optical axis, since the TE and TM components of the probe light propagate through the same SOA at the same time [113]. However, experiments show that this induced phase shift is polarization dependent [113], which implies birefringence. In addition, the absolute value of this phase inversely depends on the SOA gain level, but the pump light integral shows a positive correlation. This means that it is not well justified to attribute this phase shift to the TPA solely.

So far, the theory of ultrafast processes in a SOA is mainly focused on the carrier distribution dynamics in energy space. In an isotropic material, there is no polarization discrimination for the carrier optical transition at the same energy level. Therefore, no matter how the carriers are distributed in energy within the bands, this dynamics will not result in birefringence. Beside the distribution in energy, carrier non-equilibrium orientation takes place in momentum space, which can be understood by the electron-hole dipole polarization dynamics. The interaction of the carriers with the electromagnetic field is the scalar product of the electric field $\vec{E}$ and the transition dipole moment $e\vec{r}$. So the interaction of a cross-polarized electric field and the dipole will be different from that in the co-polarized case. When the orientation of the dipoles is randomly distributed, no macroscopic birefringence will occur. However, if the initial polarizations of the electron-hole dipoles are not randomly oriented, but a certain orientation is preferred, then the response of these dipoles will discriminate between two different polarizations of the electric field and so result in birefringence.

In the optically excited semiconductor, the pump light can lead to this ultrashort dipole polarization non-equilibrium. The pump light generates transition dipoles with preferred orientation and also active carriers $n_a^+(\hat{\rho})$ or absorb carriers $n_a^-(\hat{\rho})$ with anisotropic momentum distribution. The rapidly decaying polarization, formed by the hot dipoles with preferred polarization will be experienced by the probe light for an ultrashort time, before the polarization relaxes to equilibrium zero.

In this chapter we investigate the dipole polarization non-equilibrium in an optically exited semiconductor, its influence on the response to the polarized electric field, and the resulting ultrashort birefringence. This is done by the Semiconductor Bloch Equations (SBE). Some possible mechanisms that may account for the measured birefringence will be given.

5.1 Anisotropic momentum space filling[1]

In this section we first give the standard analysis of the birefringence caused by a linearly polarized pump pulse as it is experienced by a probe pulse. We shall argue that this analysis is incomplete and predicts too weak birefringence to account for the large polarization rotation observed in experiments [106, 108].

---

[1] Wang W., Allaart K., Lenstra D., CLEO/EQEC, Munich, Germany, Jun. 12-17, (2005)

The experimental situation on which we shall focus is that of a collinear pump and probe pulse, e.g. propagating in a bulk SOA, which are both linearly polarized. Their polarization vectors make an angle θ_0, so $\theta_0 = 0$ means co-polarization and $\theta_0 = \pi/2$ means cross-polarization. These polarization vectors, together with the direction of propagation, define the laboratory reference frame. In this frame the Bloch vector of an electron that is involved in an optical transition has polar angles θ and ϕ, as shown in Fig. 3.2, however, θ here is equivalent to ϕ of the figure, and ϕ equivalent to φ.

The major part of the optical transitions in unstrained III-V semiconductors are those between the conduction band and heavy-hole band, of which the wave function is given by [114]

$$\psi_h = -\frac{1}{\sqrt{2}} \left[(|X\rangle \cos\phi + |Y\rangle \sin\phi) \cos\theta - i(|X\rangle \sin\phi - |Y\rangle \cos\phi) - |Z\rangle \sin\theta \right] \tag{5.1}$$

with the spin parallel to the $\vec{k}$ vector. Degenerate with this is the time-reversed state. The component $|X\rangle$ denotes the normalized $p_{3/2}$ state wave function

$$|X\rangle = \sqrt{\frac{3}{8\pi}} \frac{x}{r} f_{3/2}(r) \tag{5.2}$$

and similar for $|Y\rangle$ and $|Z\rangle$. By optical transition it is connected to the conduction band state

$$|s\rangle = \sqrt{\frac{1}{4\pi}} f_s(r) \tag{5.3}$$

with the same $\vec{k}$ vector and same spin direction. A pump pulse with polarization along the z direction couples, in dipole approximation $\vec{E} \cdot \vec{r}$, only to the $|Z\rangle$ component of the heavy-hole wave function. Therefore the optical transition probability between the heavy-hole band and the conduction band is for a given direction of $\vec{k}$ proportional to $\sin^2\theta$. In other words, the excited carriers are anisotropically distributed in $\vec{k}$ space, with a $\sin^2\theta$ distribution, where θ is the angle between the polarization of the pump and the $\vec{k}$ vector of the states involved. This distribution is illustrated in Fig. 5.1. Note that it only depends on the relative angle θ between the momentum vector $\vec{k}$ and the electric polarization of the pump, irrespective of whether this is a TE or TM wave in a wave guide, or just a plane wave. Carriers excited from the light-hole band have a complementary distribution in $\vec{k}$ space, but are much smaller in number because of the much lower transition state density. They are therefore ignored in this qualitative analysis.

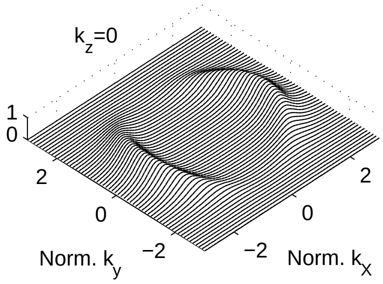

Figure 5.1: Carrier anisotropic momentum space filling.

The probe pulse, linearly polarized at an angle θ_0 relative to the pump pulse, is sensitive to a $|Z\rangle\,\mathrm{cos}\theta_0 + |X\rangle\,\mathrm{sin}\theta_0$ component of the hole wave function. The gain that it experiences due to the carriers that were excited by the pump is then, after integration over the $\sin^2\theta$ distribution, found to depend on θ_0 as

$$\text{gain} \propto N(\frac{4}{5}\mathrm{cos}^2\theta_0 + \frac{3}{5}\mathrm{sin}^2\theta_0). \tag{5.4}$$

This implies that the distribution of hot carriers produced by the pump pulse gives rise to birefringence experienced by the probe pulse. By varying the delay time between pump and probe pulse, the relaxation time with which the anisotropy of the carrier distribution in $\vec{k}$ space is dissolved can be measured.

Due to carrier-carrier and carrier-phonon collisions the anisotropic momentum distribution will decay into an isotropic one with a certain relaxation time τ_d. The hot excited carriers will also thermalize with a relaxation time τ_{SHB}, typical for spectral hole burning. Since these intra-band diffusion time scales are very short, well below one ps, the early experiments [105] could not measure these relaxation times, but only demonstrated the polarization rotation of the probe as a consequence of the birefringence. With ultrashort pulses of only 9 fs duration, Portella et al. [106] found a relaxation time of 30 - 50 fs. This is rather different from times of a few hundred fs found in theoretical studies. For instance Binder et al. [85], using both a Green function and a Boltzmann collision approach, distinguish between τ_d and τ_{SHB}, but find them to be both typically in the range 100 - 200 fs, with τ_d somewhat longer than τ_{SHB}. The discrepancy with the measured relaxation times suggests that the analysis given so far is incomplete and that a mechanism with another, shorter, time scale must be included. Another indication for this is that the anisotropy in Eq.(5.4) is rather weak and difficult to reconcile with the large polarization rotation seen in recent experiments [106]. If the latter are interpreted as due to the anisotropy in the refractive index, one expects a corresponding anisotropy in gain that is much larger than found in experiment. This observation motivated us to theoretically reconsider the anisotropy of the dynamic susceptibility up to third order (first order in the weak probe, second order in the strong pump field) starting from the basic equations of motion. We will show that a more pronounced birefringence can indeed occur when pump and probe are very close, or even overlapping, in time.

5.2 Birefringence within the "dephasing" time

5.2.1 Effective two-band model

We examine the dynamic non-linear susceptibility by calculating the polarization up to first order in the probe and second order in the pump pulse. It will turn out that the anisotropy discussed in the previous section is only part of the total anisotropy that may occur for ultrashort delay times.

The Hamiltonian for the charge carriers in interaction with a classical optical field $\vec{E}(\mathrm{t})$ is, in dipole approximation, written as

$$H = \sum_{n\vec{k},s} E_{nk} c^{\dagger}_{n\vec{k},s} c_{n\vec{k},s} - e\vec{E}(t) \cdot \sum_{n} \sum_{n'} \sum_{\vec{k},s} \langle n\vec{k}, s | \, \vec{r} \, | n'\vec{k}, s \rangle \, c^{\dagger}_{n\vec{k},s} c_{n'\vec{k},s} + V^{\mathrm{Coul}}, \tag{5.5}$$

with n and n' band indices and $E_{n,k}$ band energies. The label s indicates with $s = 1$ angular momentum parallel to the $\vec{k}$ vector, as is the case for the heavy-hole state, Eq.(5.1) or, with $s = -1$, antiparallel. The residual Coulomb interaction V^{Coul} will be represented in the following only by relaxation terms, as it does not primarily generate anisotropy, but only contributes to its decay. Thereby, and because we restrict ourselves to the heavy-hole band and conduction band only, we have effectively dynamical equations for independent two-level systems for each $\vec{k}, s$. The polarization is given by

$$\vec{P}(t) = \sum_{\vec{k},s} \langle c\vec{k}, s | \, e\vec{r} \, | h\vec{k}, s \rangle \, \rho_{\vec{k},s} + \mathrm{cc}, \tag{5.6}$$

with $\rho_{\vec{k},s}$ the expectation value of the off-diagonal element of the density operator

$$\rho_{\vec{k},s} = \langle c^{\dagger}_{c\vec{k},s} c_{h\vec{k},s} \rangle. \tag{5.7}$$

For the heavy-hole states the dipole matrix elements can be expressed as

$$\langle c\vec{k}, s | \, e\vec{r} \, | h\vec{k}, s \rangle = eR\hat{n}^{s}_{\vec{k}}, \tag{5.8}$$

with R the radial matrix element

$$R = \frac{\mathrm{e}}{\sqrt{3}} \int dr f_s^*(r) f_{3/2}(r) r \tag{5.9}$$

and $\hat{n}_{\vec{k}}^s$ a circular unit vector, right handed or left handed with respect to the direction of $\vec{k}$;

$$\hat{n}_{\vec{k}}^{\pm} = \mp(\hat{x}' \pm i\hat{y}') \tag{5.10}$$

in a (x', y', z') reference frame with $\hat{k}$ as the z' axis. Expressed in laboratory frame coordinates, cf. Eq. (5.1),

$$\hat{n}_{\vec{k}}^s = \frac{1}{\sqrt{2}}(\hat{x}\{-s\cos\phi\cos\theta + i\sin\phi\} + \hat{y}\{-s\sin\phi\cos\theta - i\cos\phi\} + s\hat{z}\sin\theta). \tag{5.11}$$

This expression is required for integrations over all directions of $\vec{k}$, while Eqs. (5.8) and (5.9) are a transparent short notation.

We further introduce the short-hand notations:

$$\vec{E}(t) \cdot \hat{n}_{\vec{k}}^s \equiv E_{\vec{k}}^s = -(E_{\vec{k}}^{-s})^*, \tag{5.12}$$

$$\langle c^\dagger_{c\vec{k},s} c_{c\vec{k},s}\rangle - \langle c^\dagger_{h\vec{k},s} c_{h\vec{k},s}\rangle \equiv N_{\vec{k},s} \tag{5.13}$$

and $\omega_k = \frac{1}{\hbar}(E_{ck} - E_{hhk})$, which is independent of s and of the direction of $\vec{k}$. The Heisenberg equations of motion, equivalent to the well known semiconductor Bloch equations [2, 7], then take the form:

$$\frac{d}{dt} N_{\vec{k}s} = 2\frac{ieR}{\hbar}(E_{\vec{k}}^s \rho_{\vec{k}s} + E_{\vec{k}}^{-s} \rho_{\vec{k}s}^*) - \frac{1}{T_1}(N_{\vec{k}s} - N_k^{eq}) \tag{5.14}$$

$$\frac{d}{dt}\rho_{\vec{k}s} = i\omega_k \rho_{\vec{k}s} - \frac{ieR}{\hbar} E_{\vec{k}}^{-s} N_{\vec{k}s} - \frac{1}{T_2}\rho_{\vec{k}s}. \tag{5.15}$$

Here we have introduced relaxation times T_1 and T_2, the meaning of which we shall discuss shortly. Note that the equilibrium population N_k^{eq}, before the arrival of the light pulses, will be isotropic in $\vec{k}$ space, also in case of electronic pumping (from a non-magnetic source). The statistical ensemble average of the polarizations, $\rho_{\vec{k}s}$, will be zero before the pulses arrive.

After solving these coupled equations, for a given classical field $\vec{E}$(t), the polarization is obtained as

$$\vec{P}(t) = eR \sum_{\vec{k},s} \{\hat{n}_{\vec{k}}^s \rho_{\vec{k}s}(t) - \hat{n}_{\vec{k}}^{-s} {\rho_{\vec{k}s}}^*(t)\}. \tag{5.16}$$

5.2.2 Phenomenological relaxation times T_1 and T_2

The relaxation time T_1 represents the return to an equilibrium population in $\vec{k}$ space. Anisotropy of the population in $\vec{k}$ space gives rise to part of the pump - probe induced birefringence discussed here. One may distinguish two aspects of the collision mechanisms by which the effect of this anisotropy decays. One is the state filling by scattering of electrons out of states with different energy. The associated time is that observed in hole burning experiments with unpolarized light, τ_{SHB}. The other aspect is that of random reorientation of the momenta by scattering events. The time associated with that *isotropic* state filling τ_d is the more relevant one for birefringence, if it is shorter than τ_{SHB}. Theoretical studies [85] indicate, however, that τ_d is somewhat larger than τ_{SHB} and that both times are roughly 100 - 200 fs. So we introduce here only one relaxation time T_1 that is supposed to include both. Of course T_1 depends on the carrier densities [2].

The polarization relaxation time T_2 is the time after which the special phase relation between the electron and ”its” hole, with the same $\vec{k}s$, is lost. That is about a scattering time. For room temperature equilibrium plasmas it is typically a few tens of fs [2]. This is comparable to a 30 - 50 fs decay time of the birefringence measured with 9 fs pulses [106]. In a later section, 5.2.5C, we shall assume that T_2 also comprises the decay of the macroscopic polarization as a consequence of dephasing.

5.2.3 Formal solution of the equations of motion

The coupled equations of motion are formally solved by time integration of Eq. (5.15) substituted in Eq. (5.32), which gives

$$\begin{aligned} N_{\vec{k}s}(t) = N_k^{eq} + \frac{2e^2R^2}{\hbar^2}\int_{-\infty}^{t} dt' \int_{-\infty}^{t'} dt''(E^s_{\vec{k}}(t')E^{-s}_{\vec{k}}(t'')\mathrm{e}^{i\omega_k(t'-t'')} \\ +E^{-s}_{\vec{k}}(t')E^s_{\vec{k}}(t'')\mathrm{e}^{-i\omega_k(t'-t'')})N_{\vec{k}s}(t'')\mathrm{e}^{(t''-t')/T_2}\mathrm{e}^{(t'-t)/T_1}, \qquad (5.17) \end{aligned}$$

which gives a perturbation expansion for $N_{\vec{k}s}$ in powers of a given electric field intensity E^2. Note that if the electric field is only that of the linearly polarized pump pulse $\vec{E}_{pump} = \hat{z}E_1(t)$, then the factor $E^s_{\vec{k}}E^{-s}_{\vec{k}}$ just gives the $\sin^2\theta$ distribution for $N_{\vec{k}s} - N_k^{eq}$ mentioned in section 5.1. One then neglects the interference term between the pump and weak probe pulse, $\vec{E}_{probe} = (\hat{z}\cos\theta_0 + \hat{x}\sin\theta_0)E_2(t)$. That interference gives an angle dependent factor

$$\sum_s \vec{E}_{pump}\cdot\hat{n}^s_{\vec{k}}\vec{E}_{probe}\cdot\hat{n}^{-s}_{\vec{k}} = -2\cos\theta_0\sin^2\theta + 2\sin\theta_0\sin\theta\cos\theta\cos\phi. \qquad (5.18)$$

So the complete population distribution becomes, for $\theta_0 \neq 0$, dependent on the azimuthal angle ϕ of $\vec{k}$. Note that this is only effective if the time delay between pump and probe pulse is less than T_2. Though this deviation from the $\sin^2\theta$ distribution is small, in the case of a weak probe pulse, its effect on the observed birefringence is not negligible, as the strong probe pulse will experience this azimuthal asymmetry. The effect on the third order polarization is still linear in the probe field, quadratic in the pump field and thererefore of comparable magnitude as the effect of the strong $\sin^2\theta$ dependence of the population.

The complete expression for the part of the third order polarization, Eq.(5.16), that is linear in the probe and quadratic in the pump pulse is obtained after integration over all directions of $\vec{k}$ and summation over s as:

$$\vec{P}^{(3)}(t) = -\frac{e^4R^4}{h^3}\frac{8\pi}{15}\int k^2 dk N_k^{eq}\int_{-\infty}^{t} dt_1 \int_{-\infty}^{t_1} dt_2 \int_{-\infty}^{t_2} dt_3 e^{(t_1-t)/T_2}$$
$$e^{(t_3-t_2)/T_2} e^{(t_2-t_1)/T_1}\Big\{ \sin\{\omega_k(t_1-t)\}\cos\{\omega_k(t_3-t_2)\}\{(4\hat{z}\cos\theta_0 + 3\hat{x}\sin\theta_0)$$
$$E_2(t_1)E_1(t_2)E_1(t_3) + (4\hat{z}\cos\theta_0 + \frac{1}{2}\hat{x}\sin\theta_0)E_1(t_1)(E_2(t_2)E_1(t_3)$$
$$+E_1(t_2)E_2(t_3)) - \frac{5}{2}\hat{x}\sin\theta_0\cos\{\omega_k(t_1-t)\}\sin\{\omega_k(t_3-t_2)\}$$
$$E_1(t_1)\{E_2(t_2)E_1(t_3) - E_1(t_2)E_2(t_3)\}\Big\}. \tag{5.19}$$

The first term has the familiar anisotropy factor 4 : 3 for co-polarized *vs.* cross-polarized probe. In the limit $T_2 \to 0$ it follows instantaneously the probe pulse and is proportional to the power of the pump pulse. The next terms exhibit a much stronger, 8 : 1, anisotropy. They reflect the interference between pump and probe, within the polarization decay time T_2. The last term is also strongly anisotropic; it vanishes in case of complete time overlap between pump and probe.

Eq.(5.19) can be expressed in a somewhat more transparent form if we assume the electric field strength to be given by

$$\vec{E}(t) = \tilde{E}_1(t)\cos(\omega t)\hat{z} + \tilde{E}_2(t)\cos(\omega t + \psi)(\cos\theta_0\hat{z} + \sin\theta_0\hat{x}), \tag{5.20}$$

with $\tilde{E}_1(t)$ and $\tilde{E}_2(t)$ slowly varying envelope functions. The full expression then is:

$$
\begin{aligned}
\vec{P}^{(3)}(t) \quad = \quad & -\frac{e^4R^4}{h^3}\frac{8\pi}{15}\int k^2 dk N_k^{eq} \int_{-\infty}^{t} dt_1 \int_{-\infty}^{t_1} dt_2 \int_{-\infty}^{t_2} dt_3 \\
& \mathrm{e}^{(t_1-t)/T_2}\mathrm{e}^{(t_3-t_2)/T_2}\mathrm{e}^{(t_2-t_1)/T_1} \\
& \big([(4\hat{z}\cos\theta_0 + 3\hat{x}\sin\theta_0)\tilde{E}_2(t_1)\tilde{E}_1(t_2)\tilde{E}_1(t_3) \\
& +(4\hat{z}\cos\theta_0 + \frac{1}{2}\hat{x}\sin\theta_0)\tilde{E}_1(t_1)\{\tilde{E}_2(t_2)\tilde{E}_1(t_3) + \tilde{E}_1(t_2)\tilde{E}_2(t_3)\}] \\
& \sin\{\omega_k(t_1-t)\}\cos\{\omega_k(t_3-t_2)\}\cos(\omega t_1+\psi)\cos(\omega t_2)\cos(\omega t_3) \\
& +\sin\psi(4\hat{z}\cos\theta_0 + \frac{1}{2}\hat{x}\sin\theta_0)\sin\{\omega_k(t_1-t)\}\cos\{\omega_k(t_3-t_2)\} \\
& [\sin\{\omega(t_1-t_2)\}\cos(\omega t_3)\tilde{E}_1(t_1)\tilde{E}_2(t_2)\tilde{E}_1(t_3) \\
& \qquad +\sin\{\omega(t_1-t_3)\}\cos(\omega t_2)\tilde{E}_1(t_1)\tilde{E}_1(t_2)\tilde{E}_2(t_3)] \\
& +\sin\psi\frac{5}{2}\hat{x}\sin\theta_0\cos\{\omega_k(t_1-t)\}\sin\{\omega_k(t_3-t_2)\} \\
& [\sin\{\omega(t_1-t_3)\}\cos(\omega t_2)\tilde{E}_1(t_1)\tilde{E}_1(t_2)\tilde{E}_2(t_3) \\
& \qquad -\sin\{\omega(t_1-t_2)\}\cos(\omega t_3)\tilde{E}_1(t_1)\tilde{E}_2(t_2)\tilde{E}_1(t_3)]\big).
\end{aligned} \tag{5.21}
$$

With the argument that the phase difference ψ will vary over the whole volume of the scatterer, the factor $\sin\psi$ vanishes on the average and one obtains

$$
\begin{aligned}
\vec{P}^{(3)}(t) \quad = \quad & -\frac{e^4R^4}{h^3}\frac{8\pi}{15}\int k^2 dk N_k^{eq} \int_{-\infty}^{t} dt_1 \int_{-\infty}^{t_1} dt_2 \int_{-\infty}^{t_2} dt_3 \\
& \mathrm{e}^{(t_1-t)/T_2}\mathrm{e}^{(t_3-t_2)/T_2}\mathrm{e}^{(t_2-t_1)/T_1} \\
& [(4\hat{z}\cos\theta_0 + 3\hat{x}\sin\theta_0)\tilde{E}_2(t_1)\tilde{E}_1(t_2)\tilde{E}_1(t_3) \\
& +(4\hat{z}\cos\theta_0 + \frac{1}{2}\hat{x}\sin\theta_0)\tilde{E}_1(t_1)\{\tilde{E}_2(t_2)\tilde{E}_1(t_3) + \tilde{E}_1(t_2)\tilde{E}_2(t_3)\}] \\
& \sin\{\omega_k(t_1-t)\}\cos\{\omega_k(t_3-t_2)\}\cos(\omega t_1+\psi)\cos(\omega t_2)\cos(\omega t_3).
\end{aligned} \tag{5.22}
$$

This expression suggests rather weak birefringence, 4:3 for co-polarized *vs.* cross polarized light during a decay time T_1 and a short birefringence, during decay time T_2 between the pulses, but with a 8:1 ratio for co- *vs.* cross-polarization. Closer inspection of the time structure of the integrals reveals, however, that the terms with the ratio 8:1 vanish if the probe $\vec{E}_2(t)$ comes after the pump $\vec{E}_1(t)$ and is well separated from it in time. To elucidate this point further, we now consider two specific cases separately. First pulses that are many oscillation periods long, say a few hundred fs and may partially overlap in time. Thereafter ultrashort pulses, not overlapping in time.

5.2.4 Adiabatic following approximation

If the pulse length is many times the oscillation period of the light, then the polarization $\vec{P}_{\vec{k},s}(t)$ will to a good approximation "adiabatically follow" the oscillation of the field strength. In the "resonant" or "rotating wave" approximation one neglects terms with factor $(\omega_k + \omega)^{-1}$ as compared to $(\omega_k - \omega)^{-1}$ and obtains in this way

$$\rho_{\vec{k},s}(t) = \frac{\mathrm{e}^{i\omega t}}{\omega_k - \omega + \frac{i}{T_2}} \frac{eR}{2\hbar}[\tilde{E}_1(t)\hat{z} + \tilde{E}_2(t)\mathrm{e}^{i\psi}(\mathrm{cos}\theta_0\hat{z} + \mathrm{sin}\theta_0\hat{x})] \cdot \hat{n}_{\vec{k}}^{-s} N_{\vec{k},s}(t). \tag{5.23}$$

Substitution in the equation for $N_{\vec{k},s}(t)$ then gives

$$\begin{aligned} N_{\vec{k}s}(t) &= N_k(t_0) - \frac{e^2R^2}{2\hbar^2}\frac{T_2}{(\omega-\omega_k)^2T_2^2+1}\int_{t_0}^{t} dt'\mathrm{e}^{\frac{1}{T_1}(t-t')}N_{\vec{k}s}(t') \\ &\quad \big(\tilde{E}_1^2(t')\mathrm{sin}^2\theta + 2\tilde{E}_1(t')\tilde{E}_2(t')\{\mathrm{cos}\psi(\mathrm{cos}\theta_0\mathrm{sin}^2\theta - \sin\theta_0\mathrm{sin}\theta\mathrm{cos}\theta\mathrm{cos}\phi) \\ &\quad +s\mathrm{sin}\psi\mathrm{sin}\theta_0\mathrm{sin}\theta\mathrm{sin}\phi\}\big). \end{aligned} \tag{5.24}$$

This can be read as a perturbation series for the populations, as long as the field strengths do not change. The s dependence of the interference term is transparent: if $\mathrm{sin}\psi \neq 0$, for instance $\psi = \frac{\pi}{2}$, the combination of pump and probe results in elliptically polarized light. A state with polar coordinates θ, ϕ of $\vec{k}$ is then most sensitive to this left- or right-handed elliptically polarized light; depending on the sign of s a $m' = 3/2$ to $m' = 1/2$ or a $m' = -3/2$ to $m' = -1/2$ transition.

Inserting this expression for $N_{\vec{k}s}$ in the expression for the polarization and applying again the "adiabatic following" approximation one obtains the terms of order $E_2E_1^2$

$$\begin{aligned} \vec{P}^{(3)}(t) &= -\frac{e^4R^4}{h^3}\frac{2\pi}{15}\int k^2dkN_k^{eq}\frac{T_2}{\{(\omega_k-\omega)^2T_2^2+1\}^{3/2}} \\ &\quad [\{\tilde{E}_2(t)(4\hat{z}\mathrm{cos}\theta_0 + 3\hat{x}\mathrm{sin}\theta_0)\mathrm{cos}(\omega t + \psi - \alpha_k)\} \\ &\qquad \int_{-\infty}^{t} dt'\tilde{E}_1^2(t')\mathrm{e}^{-\frac{1}{T_1}(t-t')} \\ &\quad +\tilde{E}_1(t)\{(8\hat{z}\mathrm{cos}\theta_0 + \hat{x}\mathrm{sin}\theta_0)\mathrm{cos}(\omega t + \psi - \alpha_k) \\ &\qquad +(8\hat{z}\mathrm{cos}\theta_o - 4\hat{x}\mathrm{sin}\theta_0)\mathrm{sin}\psi\mathrm{sin}(\omega t - \alpha_k)\} \\ &\qquad \int_{-\infty}^{t} dt'\tilde{E}_1(t')\tilde{E}_2(t')\mathrm{e}^{-\frac{1}{T_1}(t-t')}]. \end{aligned} \tag{5.25}$$

The angle α_k is introduced here as

$$\cos\alpha_k = \frac{\omega_k - \omega}{\{(\omega_k-\omega)^2 + T_2^{-2}\}^{\frac{1}{2}}}; \quad \sin\alpha_k = \frac{T_2^{-1}}{\{(\omega_k-\omega)^2 + T_2^{-2}\}^{\frac{1}{2}}}. \tag{5.26}$$

For comparison we give the linear polarization of the probe:

$$\begin{aligned}\vec{P}^{(2)}(t) &= -\frac{e^2R^2}{3\pi h}\int k^2 dk N_k^{eq}\frac{T_2}{\{(\omega_k-\omega)^2T_2^2+1\}^{1/2}} \\ &\quad [\tilde{E}_2(t)(\hat{z}\text{cos}\theta_0 + \hat{x}\text{sin}\theta_0)\text{cos}(\omega t + \psi - \alpha_k)]. \end{aligned}\tag{5.27}$$

Again one notes that the last term of $\vec{P}^{(3)}$ contributes only if the pump $\vec{E}_1$ and the probe $\vec{E}_2$ overlap in time. If not, one has again simply the anisotropy factor 4:3 for co-polarized vs. cross-polarized light. The other terms seem to show a more pronounced anisotropy. The scalar product $\vec{E}_2(t)\cdot\vec{P}^{(3)}(t)$ shows, however, that these terms together yield also a 4:3 anisotropy of the gain and refractive index, but they do give a doubling of the birefringence during the overlap of the two pulses in time. The decay time T_2 does not contribute to this time overlap here, as a consequence of the adiabatic following and resonant approximation.

5.2.5 Non-overlapping ultrashort linearly polarized pulses

5.2.5A. Polarization and excitation by a short pump pulse

In the adiabatic following approximation, Eq.(5.23), the polarization is zero at the time between the pulses, if these are well separated in time. The more exact expression, with the upper label indicating that it is due to the pump pulse $\vec{E}_1(t)$ only, after passage of a short pump pulse is, by integration of Eq.(5.15):

$$\rho_{\vec{k}s}^{(1)}(t) = A_{\vec{k}s}^{(1)}\mathrm{e}^{i\omega_k(t-t_p)}\mathrm{e}^{-(t-t_p)/T_2} \tag{5.28}$$

with t_p the time around which the pump pulse is centered and amplitude

$$A_{\vec{k}s}^{(1)} = \frac{-ieR}{\hbar}\int dt\vec{E}_1(t)\cdot\hat{n}_{\vec{k}}^{-s}N_{\vec{k}s}(t)\mathrm{e}^{(i\omega_k-\frac{1}{T_2})(t_p-t)}. \tag{5.29}$$

This amplitude reflects an effective impulse given to the transition dipole oscillator by the electric force field $\vec{E}_1(t)$ during the passage of the pump pulse. The contribution of each $\vec{k}s$ to the total polarization Eq.(5.16) is

$$\begin{aligned}\vec{P}_{\vec{k}s}^{(1)}(t) &= e\vec{r}_{\vec{k}s}^{(1)}(t) \\ &= eR\left(\hat{n}_{\vec{k}}^{s}A_{\vec{k}s}^{(1)}\mathrm{e}^{i\omega_k(t-t_p)} - \hat{n}_{\vec{k}}^{-s}A_{\vec{k}s}^{(1)*}\mathrm{e}^{-i\omega_k(t-t_p)}\right)\mathrm{e}^{-(t-t_p)/T_2}. \end{aligned}\tag{5.30}$$

The transition dipole vector $\vec{r}^{(1)}_{\vec{k}s}(t)$ obviously satisfies the equation for a damped harmonic oscillator

$$\left(\frac{d^2}{dt^2} + \omega_k^2 + \frac{1}{T_2^2} + \frac{2}{T_2}\frac{d}{dt}\right)\vec{r}^{(1)}_{\vec{k}s}(t) = 0. \tag{5.31}$$

So the probe $\vec{E}_2(t)$, coming after the pump pulse $\vec{E}_1(t)$, sees an assembly of oscillating dipoles, each with its own frequency ω_k.

It also sees a distribution of excited carriers

$$N^{(1)}_{\vec{k}s}(t) = N_k^{eq} + \sin^2\theta\Delta^{(1)}_{\vec{k}s}\mathrm{e}^{-(t-t_p)/T_1} \tag{5.32}$$

with

$$\begin{aligned}\Delta^{(1)}_{\vec{k}s}(t) = \quad & -\frac{4e^2R^2}{\hbar^2}\int_{-\infty}^{t} dt' \int_{-\infty}^{t'} dt'' E_1(t')E_1(t'') \\ & \cos\{\omega_k(t'-t'')\}N^{(1)}_{\vec{k}s}(t'')e^{(t''-t')/T_2}e^{(t'-t_p)/T_1}.\end{aligned} \tag{5.33}$$

By substitution of Eq.(5.32) in Eq.(5.33) one obtains a power series in $\sin^2\theta$ and at the same time in the electric field strength:

$$N^{(1)}_{\vec{k}s}(t) = N_k^{eq} + \sin^2\theta\Delta'^{(1)}_{\vec{k}s}e^{-(t-t_p)/T_1} + \sin^4\theta\Delta''^{(1)}_{\vec{k}s}e^{-2(t-t_p)/T_1} + \ldots \tag{5.34}$$

Another way [115] to analyze the carrier population distribution Eq.(5.33) is to write the pump pulse as a product of slowly varying envelop function and the rapidly varying carrier wave: $\vec{E}_1(t) = \hat{z}\tilde{E}_1(t)\cos(\omega t)$ and use the Taylor expansion for the envelope function

$$\tilde{E}_1(t'')N^{(1)}_{\vec{k}s}(t'') = \tilde{E}_1(t')N^{(1)}_{\vec{k}s}(t') + (t''-t')\frac{d}{dt'}\left(\tilde{E}_1(t')N^{(1)}_{\vec{k}s}(t')\right) + \ldots \tag{5.35}$$

which is only meaningful, however, for small time difference $t''-t'$, during the pulse. One then obtains for the part quadratic in $\tilde{E}_1$

$$\begin{aligned}\Delta'^{(1)}_{\vec{k}s}(t) \quad = \quad & -\frac{2e^2R^2}{\hbar^2}\Big[\frac{2T_2}{(\omega-\omega_k)^2T_2^2+1}\int_{-\infty}^{t} dt'|\tilde{E}_1(t')|^2N^{(1)}_{\vec{k}s}(t')\mathrm{e}^{-(t-t_p)/T_1} \\ & + \frac{\{(\omega-\omega_k)^2T_2^2-1\}T_2^2}{\{(\omega-\omega_k)^2T_2^2+1\}^2}|\tilde{E}_1(t)|^2N^{(1)}_{\vec{k}s}(t)\Big].\end{aligned} \tag{5.36}$$

The first term tells that the effect on the population is proportional to the power of the pump pulse, with efficiency factor $N^{(1)}_{\vec{k}s}$ and has a resonance structure. The second term describes the phenomenon of the "light induced hole" [115], due to its typical resonance form. This term is instantaneous and therefore of no concern for us here. It should be noted that the importance of the next terms in the expansion is not clear. The higher derivatives in Eq.(5.35) will gain significance as pulses become shorter.

5.2.5B. Polarization during probe pulse

The polarization during the passage of the probe pulse is again composed of the contributions of the transition dipoles for all $\vec{k}s$. If the electric field strength is given as

$$\vec{E}_2(t) = (\hat{z}\cos\theta_0 + \hat{x}\sin\theta_0)E_2(t) \tag{5.37}$$

then the equation of motion for the transition dipole becomes

$$\begin{aligned}
[\frac{d^2}{dt^2} &+ \omega_k^2 + \frac{1}{T_2^2} + \frac{2}{T_2}\frac{d}{dt}]\vec{r}_{\vec{k}s}(t) = \\
&- \frac{eR^2}{\hbar}(\frac{1}{T_2} + \frac{d}{dt})\left(E_2(t)N_{\vec{k}s}(t)\right) \\
&\quad s[\cos\theta_0(\hat{x}\sin\theta\sin\phi - \hat{y}\sin\theta\mathrm{cox}\phi) + \sin\theta_0(\hat{y}\cos\theta\cos\phi - \hat{z}\cos\theta\sin\phi)] \\
&+ \frac{eR^2}{\hbar}\omega_k E(t)N_{\vec{k}s}(t) \\
&\quad [\cos\theta_0\{(\hat{x}\cos\phi + \hat{y}\sin\phi)\cos\theta - \hat{z}\sin\theta\} \\
&\quad +\sin\theta_0\{\hat{x}(\cos^2\theta\cos^2\phi + \sin^2\phi) + \hat{y}\sin\phi\cos\phi(\cos^2\theta + 1) \\
&\qquad +\hat{z}\sin\theta\cos\theta\cos\phi\}].
\end{aligned} \tag{5.38}$$

The complicated form of this driving term, which is not parallel to the driving electric field $\vec{E}_2$, is due to the fact that the transition dipole between heavy-hole and conduction band (for isotropic semiconductor) can only oscillate in a direction perpendicular to the momentum vector $\vec{k}$. The total polarization of all dipoles together does oscillate in the direction of the force field $\vec{E}$, provided that there was no preceding pulse with a different polarization direction. For

$\vec{P} = e\sum_{\vec{k}s}\vec{r}_{\vec{k}s} = \sum_k\sum_s\int d\Omega_{\vec{k}}$, we obtain, with substitution of Eq. (5.34)

$$\begin{aligned}
[\frac{d^2}{dt^2} &+ \omega_k^2 + \frac{1}{T_2^2} + \frac{2}{T_2}\frac{d}{dt}]\vec{P}(t) \\
&= \frac{e^2R^2}{\hbar}\sum_k \omega_k E_2(t) \\
&\quad [\frac{16\pi}{3}N_k^{eq}(\hat{z}\cos\theta_0 + \hat{x}\sin\theta_0) + \Delta_k^{'(1)}\mathrm{e}^{-(t-t_p)/T_1}\frac{8\pi}{75}(4\hat{z}\cos\theta_0 + 3\hat{x}\sin\theta_0) \\
&\quad + \Delta_k^{''(1)}\mathrm{e}^{-2(t-t_p)/T_1}\frac{64\pi}{105}(3\hat{z}\cos\theta_0 + 2\hat{x}\sin\theta_0) + \ldots] \qquad (5.39)
\end{aligned}$$

This series expansion, of which the second term within the brackets is quadratic in the pump field $\vec{E}_1$, see Eq. (5.36), shows the ratio 4:3 birefringence, the last term and those of still higher order in the pump field show the successively more pronounced ratios 3:2, 8:5, ... These terms have also successively shorter decay times $T_1, \frac{1}{2}T_1, \frac{1}{3}T_1$, etc. We do not consider this as the explanation of the ultrafast relaxation of the birefringence observed by Portella et al. [106], because there are at least two more effects that have to be considered.

5.2.5C. Effect of pump polarization on propagating probe field strength

In the derivations given so far, we have assumed that the probe field has a given polarization direction $\hat{z}\cos\theta_0 + \hat{x}\sin\theta_0$. This assumption is not completely adequate, because the propagating probe field is influenced by the polarization that was produced by the pump pulse Eq.(5.30), according to the well known equation [5] for the envelope functions (with $\hat{y}$ the propagation direction)

$$\frac{d}{dy}\tilde{E}(y,t) \propto \tilde{P}(y,t) \qquad (5.40)$$

where the right hand side is to a good approximation the macroscopic polarization, the sum of contributions Eq.(5.30), caused by the pump pulse. This has a very short decay time T_2 and the direction of the pump pulse polarization $\hat{z}$. Therefore, without doing an explicit calculation, one can conclude that the factors in front of $\hat{z}\cos\theta_0$ in Eq.(5.39) are modified during a typical decay time of the polarization that was produced by the pump. We consider this polarization once more:

$$\begin{aligned}\vec{P}^{(1)}(t) &= eR\sum_{\vec{k}s}\left(\hat{n}^{s}_{\vec{k}}A^{(1)}_{\vec{k}s}\mathrm{e}^{i\omega_k(t-t_p)} - \hat{n}^{-s}_{\vec{k}}A^{(1)*}_{\vec{k}s}\mathrm{e}^{-i\omega_k(t-t_p)}\right)\mathrm{e}^{(t-t_p)/T_2} \\ &= -\hat{z}\frac{16\pi e^2R^2}{3\hbar}\sum_{k}N_k\int^{t}dt'E_1(t')\mathrm{sin}\{\omega_k(t-t')\}\mathrm{e}^{(-t-t')/T_2}\end{aligned} \tag{5.41}$$

One notes that this polarization is a superposition of terms with different oscillation frequencies ω_k and already thereby it will decay faster as the pulse was shorter, exciting a broader frequency range. For a pulse of 9fs this "mutual running out of phase" will give a loss of macroscopic polarization in about 30 fs. One may then assume that this depolarization is roughly incorporated in a decay time T_2 of such a magnitude.

5.2.5D. Contribution of the light hole band and case of strain

It is well known that the charge carriers that are excited from the light-hole band will have a distribution in momentum space that is complementary to that of carriers excited from the heavy-hole band [85, 105]. If both bands would contribute equally, the distribution would be isotropic in $\vec{k}$ space. So there would be no $\sin^2\theta$ term in Eq.(5.32), no discrimination between the $\hat{z}$ and $\hat{x}$ components in Eq.(5.38) and consequently no birefringence with decay time T_1, the relaxation time of the distributions in momentum space. In unstrained bulk semiconductors this situation does not occur, but in case of strain, like in wave guides or quantum wells, it could happen. This might actually offer an interesting opportunity. For if there is no component of the birefringence with a decay time of T_1 ≈100 - 200fs, then the birefringence switch could possibly operate with a much faster relaxation time T_2, of only a few tens of fs. We do not repeat the whole analysis here for hole states in case of strain, as in chapter 3, because this would result in rather clumsy equations. Instead we only globally explore it by noting that the restriction on the oscillation direction of the transition dipole $\vec{r}_{\vec{k}s}$, *viz.* orthogonal to $\vec{k}$, no longer holds. If the direction of oscillation is free to follow that of the electric force field, the situation is similar to a classical oscillator, apart from saturation effects.

We therefore conclude this section by considering a classical, isotropic, damped harmonic oscillator model. Its equation of motion in a probe field

$E_2(\hat{z}\mathrm{cos}\theta_0 + \hat{x}\mathrm{sin}\omega_0)\mathrm{cos}(\omega t + \phi)$ is

$$\left(\frac{d^2}{dt^2} + \omega_k^2 + \frac{1}{T_2^2} + \frac{2}{T_2}\frac{d}{dt}\right)\vec{r}^{(2)}_{\vec{k}s}(t) = \frac{e\vec{E}_2}{m}\mathrm{cos}(\omega t + \phi), \tag{5.42}$$

with the solution

$$\vec{r}^{(2)}_{\vec{k}s}(t) = \frac{e\vec{E}_2}{m}[\{(\omega^2 - \omega_k^2)T_2^2 - 1\}^2 + 4\omega^2 T_2^2]^{-\frac{1}{2}}\mathrm{cos}(\omega t + \psi) \tag{5.43}$$

and

$$\mathrm{tg}(\psi - \phi) = \frac{2\omega T_2}{(\omega^2 - \omega_k^2) - 1}. \tag{5.44}$$

Note that also $\vec{r}^{(1)}_{\vec{k}s}(t)$, Eq. (5.30), satisfies the homogeneous equation Eq.(5.31), so the general solution during the passage of the probe is a superposition of $\vec{r}^{(2)}_{\vec{k}s}(t)$ and the exponentially decaying $\vec{r}^{(1)}_{\vec{k}s}(t)$. When the oscillator is not quite harmonic, there will be a small extra force term that is proportional to $\{\vec{r}^{(2)}_{\vec{k}s}(t) + \vec{r}^{(1)}_{\vec{k}s}(t)\}^3$. Consequently an extra amplitude $\vec{r}^{(3)}_{\vec{k}s}(t)$ is generated, that satisfies the equation, up to second order in the pump and first order in the probe,

$$\begin{aligned}[\frac{d^2}{dt^2} + \omega_k^2 \quad &+ \quad \frac{1}{T_2^2} + \frac{2}{T_2}\frac{d}{dt}]\vec{r}^{(3)}_{\vec{k}s}(t) \\ &\propto \quad \frac{e\vec{E}_2}{2m}\mathrm{cos}(\omega t + \psi)|A^{(1)}_{\vec{k}s}|^2 \mathrm{e}^{-2(t-t_p)/T_2}\{3\hat{z}\mathrm{cos}\theta_0 + \hat{x}\mathrm{sin}\theta_0\}\end{aligned} \tag{5.45}$$

So now the driving term for the third order polarization contains the strongly anisotropic factor $3\hat{z}\mathrm{cos}\theta_0 + \hat{x}\mathrm{sin}\theta_0$. It has a relaxation time $\frac{1}{2}T_2$. The difference with the derivations in the previous sections is that we have here an isotropic oscillator, while before only oscillation in a plane orthogonal to $\vec{k}$ was possible.

5.3 Conclusion

We have pointed out various mechanisms that can give rise to ultrafast birefringence with time constants varying from $T_1 \approx 100$ - 200 fs to $T_2 \approx 30$ fs or even $\frac{1}{2}T_2$. An all-inclusive numerical calculation which comprises all these aspects in detail is outside the scope of our present work.

Chapter 6

Polarization dependent FWM quantum beats

6.1 Introduction

Polarization-dependent four wave mixing (FWM) quantum beats [116–126] have been observed after simultaneous excitations of two optical transitions, associated with heavy-hole and light-hole. The signal magnitude and its beat phase depend on the polarization of the pump and probe with respect to each other. This phenomenon has been analyzed by applying semiconductor Bloch equations (SBE) for excitations in a six-band model with broad spectrum pump-probe pulses. [116, 117] However this theory predicts, in contradiction with experimental observations, *identical* FWM intensities for the two polarization configurations: pump and probe have either parallel or perpendicular linear (circular) polarizations. Since then, a great effort has been devoted to theoretically explain the phenomena. A successful explanation is claimed to be given by the bi-exciton theory. [118, 120–126] However, it has remained obscure why the optical Bloch equations (OBE) can describe such phenomena well in atomic optics, but the SBE fail for a similar phenomenon in a semiconductor. This interesting question should not be simply answered by the argument that the SBE formalism does not include exciton-exciton interaction processes or that a Hartree-Fock approximation is not capable to describe a quasi-stationary state of the excited semiconductor. [127]

The SBE are developed on a similar basis as the OBE. [128] The main extra efforts of the former theory are its focusing on the electron many-body effects, such as semiconductor band gap renormalization, excitonic effects, phase-space filling, etc, which do not occur in OBE. In the present work we shall completely neglect the Coulomb interaction between the charge carriers, which means that no exciton-exciton Coulomb interaction is introduced. We do not consider a specific model space of one- and two-exciton states, but just solve the Heisenberg equations of motion for the dipole to third order in the light-matter interaction. This more general approach leads to analytical expressions which show in a transparent way why the case of parallel polarization of pump and test pulse could be described without introducing the bi-exciton space, while this was not the case for cross-polarized pulses. We obtain the correct phases and relative magnitudes of the quantum beats for both cases. We therefore argue that FWM quantum beats are a purely coherent light-matter interaction effect, not fully described in the standard SBE, rather than due to electron-electron interaction processes.

We assume the electronic state of semiconductors can be completely described within a space with six states: $|1/2, \pm 1/2\rangle$, $|3/2, \pm 1/2\rangle$, and $|3/2, \pm 3/2\rangle$. The spatial properties of these states are defined in chapter 3, Tab.(3.1)

The band structure, without split-off bands, of III-V semiconductors are denoted as $|c1\rangle$, $|c2\rangle$, $|h1\rangle$, $|h2\rangle$, $|l1\rangle$, and $|l2\rangle$. Here c, h and l means conduction, heavy-hole, and light-hole bands, respectively; They are twofold degenerate by the electron spin. For small $\vec{k}$, we approximate the eigenstates of $|c1\rangle$, $|c2\rangle$ by just the basis states $|1/2, \pm 1/2\rangle$ for conduction bands, but for valence bands, the wave functions are complicated and given in chapter 3, Eq.(3.30).

6.2 Model

6.2.1 Dipole matrix elements

The light-matter interaction is described in dipole approximation as $-e\vec{\mathcal{E}}(t) \cdot \vec{r}$. We consider linearly polarized laser light in the $x-y$ plane, so the relevant components of $\vec{r}$ are $\vec{r}_\perp = \hat{x}x + \hat{y}y$. The optical transitions between the conduction bands and valence bands can be written in units of M as

$$\begin{aligned}
\langle c1| \vec{r}_\perp |h1\rangle &= u(\hat{x} + i\hat{y})(-b) \\
\langle c1| \vec{r}_\perp |h2\rangle &= u(\hat{x} + i\hat{y})(-c) + w(\hat{x} - i\hat{y})(R_h) \\
\langle c2| \vec{r}_\perp |h1\rangle &= (-w)(\hat{x} + i\hat{y})(R_h) + (-u)(\hat{x} - i\hat{y})(-c^*) \\
\langle c2| \vec{r}_\perp |h2\rangle &= (-u)(\hat{x} - i\hat{y})(b^*)
\end{aligned}$$

and

$$\begin{aligned}
\langle c1|\, \vec{r}_\perp \,|l1\rangle &= u(\hat{x}+i\hat{y})(R_l) + w(\hat{x}-i\hat{y})(c^*) \\
\langle c1|\, \vec{r}_\perp \,|l2\rangle &= w(\hat{x}-i\hat{y})(-b) \\
\langle c2|\, \vec{r}_\perp \,|l1\rangle &= (-w)(\hat{x}+i\hat{y})(b^*) \\
\langle c2|\, \vec{r}_\perp \,|l2\rangle &= (-w)(\hat{x}+i\hat{y})(c) + (-u)(\hat{x}-i\hat{y})(R_l) \qquad (6.1)
\end{aligned}$$

with $u = -\sqrt{1/2}$ and $w = \sqrt{1/6}$.

The x component of the dipole operator, that couples to x-polarized light, is expressed as a superposition of particle-hole operators with coefficients given by Eq.(6.1) as

$$x = \sum_{i,j=1,2} \sum_{\vec{k}} \left[a^\dagger_{ci,\vec{k}} a_{hj,\vec{k}} \langle ci,\vec{k}|\, x \,|hj,\vec{k}\rangle + a^\dagger_{ci,\vec{k}} a_{lj,\vec{k}} \langle ci,\vec{k}|\, x \,|lj,\vec{k}\rangle + h.c. \right]. \quad (6.2)$$

Here, the integration over the polar angles θ and ϕ of the Bloch vector $\vec{k}$ gives the total transition magnitude from electron valence states $|h_j\rangle$ and $|l_j\rangle$ to electron conduction states $|c_i\rangle$; explicitly:

$$x = \sum_{\vec{k}} \Big[a^\dagger_{c1} a_{h1}(-bu) + a^\dagger_{c1} a_{h2}(wR_h - cu) + a^\dagger_{c1} a_{l1}(uR_l + wc^*) + a^\dagger_{c1} a_{l2}(-wb) +$$
$$a^\dagger_{c2} a_{h1}(uc^* - wR_h) + a^\dagger_{c2} a_{h2}(-ub^*) + a^\dagger_{c2} a_{l1}(-wb^*) - a^\dagger_{c2} a_{l2}(wc + uR_l) + h.c. \Big]. \quad (6.3)$$

Similarly for the y-component that couples to y-polarized light:

$$y = \sum_{\vec{k}} i \Big[a^\dagger_{c1} a_{h1}(-bu) - a^\dagger_{c1} a_{h2}(wR_h + cu) + a^\dagger_{c1} a_{l1}(uR_l - wc^*) + a^\dagger_{c1} a_{l2}(wb) -$$
$$a^\dagger_{c2} a_{h1}(wR_h + uc^*) + a^\dagger_{c2} a_{h2}(ub^*) + a^\dagger_{c2} a_{l1}(-wb^*) + a^\dagger_{c2} a_{l2}(-wc + uR_l) + h.c. \Big]. \quad (6.4)$$

6.2.2 Four-wave mixing

In four-wave mixing experiments the incident light fields are a pump field $\vec{\mathcal{E}}_p$ and a much weaker probe (testing) field $\vec{\mathcal{E}}_t$, so that

$$\vec{\mathcal{E}}(t) = \vec{\mathcal{E}}_p(t) \exp(iq_p \cdot r) + \vec{\mathcal{E}}_t(t) \exp(iq_t \cdot r) + c.c. \quad (6.5)$$

For the simplicity and to still expose the essential physics, we adopt the ultrashort pulse limit [117], $\vec{\mathcal{E}}_p(t) = \vec{\mathcal{E}}_p\delta(t)$ and $\vec{\mathcal{E}}_t(t) = \vec{\mathcal{E}}_t\delta(t+\tau)$, where τ is the time delay between the incident pulses and we assume $\tau > 0$; so the test pulse preceding the pump.

We consider two cases with different pump-probe polarization configurations, case 1: both pump pulse $\vec{\mathcal{E}}_p$ and testing pulse $\vec{\mathcal{E}}_t$ are linearly polarized in x direction; case 2: the pump pulse is x polarized but the test pulse is y polarized. The initial state of the system is assumed to be in the ground state with all polarizations equal to zero, which means no specific requirements on the initial state preparation.

The detected intensity of the FWM signal is [117]

$$I_{\text{FWM}} \propto \int dt |\vec{\mathcal{P}}^{(3)}(t)|^2. \tag{6.6}$$

Here $\vec{\mathcal{P}}^{(3)}$ is the component of the third order polarization of second order in $\vec{\mathcal{E}}_p$ and first order in $\vec{\mathcal{E}}_t$:

$$\vec{\mathcal{P}}_{FWM} \propto \exp[i(2\vec{q}_p - \vec{q}_t)\cdot\vec{r}]. \tag{6.7}$$

We assume that the space of states for the carriers are the six bands described before and the Hamiltonian is the independent particle part

$$H_0 = \sum_{i,\vec{k}} \left(\epsilon_{ci\vec{k}} a^\dagger_{ci\vec{k}} a_{ci\vec{k}} + \epsilon_{hi\vec{k}} a^\dagger_{hi\vec{k}} a_{hi\vec{k}} + \epsilon_{li\vec{k}} a^\dagger_{li\vec{k}} a_{li\vec{k}} \right) \tag{6.8}$$

plus the interaction of the carriers with the light field

$$H_I = -e\left[\mathcal{E}_x(t)\cdot x + \mathcal{E}_y(t)\cdot y\right], \tag{6.9}$$

with x and y the operators of Eqs.(6.3) and (6.4), while the residual Coulomb interaction of the carriers will be represented by a relaxation parameter λ. The polarization to be calculated is the expectation value of

$$\vec{\mathcal{P}} = \hat{x}x + \hat{y}y, \tag{6.10}$$

which is obtained by applying the Heisenberg equations of motion

$$\frac{d}{dt}\vec{\mathcal{P}} = \frac{1}{i\hbar}[\vec{\mathcal{P}}, H_0 + H_I]. \tag{6.11}$$

To illustrate our point we split the operators x and y in a part that involves the heavy-hole bands and a part that involves the light-hole bands. For instance, we have

$$x = \sum_{\vec{k}} \left(X^{\dagger}_{h\vec{k}} + X_{h\vec{k}} + X^{\dagger}_{l\vec{k}} + X_{l\vec{k}} \right), \tag{6.12}$$

$$y = \sum_{\vec{k}} \left(Y^{\dagger}_{h\vec{k}} + Y_{h\vec{k}} + Y^{\dagger}_{l\vec{k}} + Y_{l\vec{k}} \right), \tag{6.13}$$

with

$$\begin{aligned} X_{h\vec{k}} &= \sum_{ij} \langle hi,\vec{k}|\, x \,|cj,\vec{k}\rangle \, a^{\dagger}_{hi,\vec{k}} a_{cj,\vec{k}} \\ X_{l\vec{k}} &= \sum_{ij} \langle li,\vec{k}|\, x \,|cj,\vec{k}\rangle \, a^{\dagger}_{li,\vec{k}} a_{cj,\vec{k}} \\ Y_{h\vec{k}} &= \sum_{ij} \langle hi,\vec{k}|\, y \,|cj,\vec{k}\rangle \, a^{\dagger}_{hi,\vec{k}} a_{cj,\vec{k}} \\ Y_{l\vec{k}} &= \sum_{ij} \langle li,\vec{k}|\, y \,|cj,\vec{k}\rangle \, a^{\dagger}_{li,\vec{k}} a_{cj,\vec{k}}. \end{aligned} \tag{6.14}$$

We shall first consider the case of cross-polarization, since this was the most troublesome to be described correctly in most previous works. For clarity of the derivation we let the pump work at the times $t = 0^+$ and at $t = 0^-$. We calculate the y-component of the polarization at $t > 0^+$ by applying the Heisenberg equation of motion for $\mathcal{P}_y$, Eq. (6.13), and introduce the notation

$$\omega_{hk} = (\varepsilon_{c\vec{k}} - \varepsilon_{h\vec{k}})/\hbar \quad \text{and} \quad \omega_{lk} = (\varepsilon_{c\vec{k}} - \varepsilon_{l\vec{k}})/\hbar. \tag{6.15}$$

In the following derivation we suppress the wave vector $\vec{k}$ in the notation, because all processes are $\vec{k}$ conserving (vertical transitions in band diagram). We now first obtain the equation

$$\begin{aligned} \frac{d}{dt} Y_h(t) &= (-i\omega_h - \lambda_2) Y_h(t) + \frac{ie}{\hbar} \mathcal{E}_p \delta(t_0^+) \sum_{ijr} \langle hi|\, y \,|cj\rangle \Big[\langle cj|\, x \,|hr\rangle \, a^{\dagger}_{hi} a_{hr} \\ & \quad - \langle cr|\, x \,|hi\rangle \, a^{\dagger}_{cr} a_{cj} + \langle cj|\, x \,|lr\rangle \, a^{\dagger}_{hi} a_{lr} \Big], \end{aligned} \tag{6.16}$$

with the solution

$$\begin{aligned} Y_h(t > 0) &= \frac{ie\mathcal{E}_p}{\hbar} \sum_{ijr} \langle hi|\, y \,|cj\rangle \Big[\langle cj|\, x \,|hr\rangle \, a^{\dagger}_{hi} a_{hr}(0^+) - \langle cr|\, x \,|hi\rangle \cdot \\ & \quad a^{\dagger}_{cr} a_{cj}(0^+) + \langle cj|\, x \,|lr\rangle \, a^{\dagger}_{hi} a_{lr}(0^+) \Big] \mathrm{e}^{-i\omega_h t} \mathrm{e}^{-\lambda_2 t}. \end{aligned} \tag{6.17}$$

Here we have introduced a relaxation or dephasing constant λ_2. The operators in the first line represent a virtual two-step process in which an electron is excited by an x-polarized field and de-excited by a y-polarized field, ending up in the same band. The operator in the last line describes a similar process, but now the electron is effectively moved, in two steps, from the light-hole to the heavy-hole band. Operators of this type are normally ignored in the SBE, as the process is far off-resonant. However, as an intermediate virtual process it does give an important contribution to the FWM signal, as we shall see in the following.

As we now require the particle-hole operators at time 0^+ we solve again the equations of motion, now with the pump at $t = 0^-$ and obtain

$$a^\dagger_{hi}a_{hr}(0^+) = \frac{ie\mathcal{E}_p}{\hbar}\sum_q \Big[\langle hr|\, x\, |cq\rangle\, a^\dagger_{hi}a_{cq}(0^-) - \langle cq|\, x\, |hi\rangle\, a^\dagger_{cq}a_{hr}(0^-)\Big], \quad (6.18)$$

$$a^\dagger_{hi}a_{lr}(0^+) = \frac{ie\mathcal{E}_p}{\hbar}\sum_q \Big[\langle lr|\, x\, |cq\rangle\, a^\dagger_{hi}a_{cq}(0^-) - \langle cq|\, x\, |hi\rangle\, a^\dagger_{cq}a_{lr}(0^-)\Big], \quad (6.19)$$

and

$$\begin{aligned} a^\dagger_{cr}a_{cj}(0^+) \quad = \quad & \frac{ie\mathcal{E}_p}{\hbar}\sum_q \Big[\langle cj|\, x\, |hq\rangle\, a^\dagger_{cr}a_{hq}(0^-) - \langle hq|\, x\, |cr\rangle\, a^\dagger_{hq}a_{cj}(0^-) + \\ & \langle cj|\, x\, |lq\rangle\, a^\dagger_{cr}a_{lq}(0^-) - \langle lq|\, x\, |cr\rangle\, a^\dagger_{lq}a_{cj}(0^-)\Big]. \end{aligned} \quad (6.20)$$

Finally we need the particle-hole operators at time 0^-, which have non-zero expectation values due to the action of the test pulse at the time $t = -\tau$. We have for instance

$$\frac{d}{dt}a^\dagger_{hi}a_{cq} = (-i\omega_h - \lambda_1)a^\dagger_{hi}a_{cq} - \frac{ie\mathcal{E}_{ty}}{\hbar}\delta(t+\tau)\big[\, \langle cq|\, y\, |hi\rangle\, a^\dagger_{cq}a_{cq} - \langle cq|\, y\, |hi\rangle\, a^\dagger_{hi}a_{hi}\big]. \quad (6.21)$$

Supposing that there were no specific correlations before the pulses arrived we then obtain, also in a similar way for the other operators:

$$\begin{aligned} a^\dagger_{hi}a_{cq}(0^-) \quad &= \quad \frac{ie\mathcal{E}_{ty}}{\hbar}\mathrm{e}^{(-i\omega_h-\lambda_1)\tau}\, \langle cq|\, y\, |hi\rangle\, [a^\dagger_{hi}a_{hi}(-\tau) - a^\dagger_{cq}a_{cq}(-\tau)] \\ a^\dagger_{cq}a_{hr}(0^-) \quad &= \quad \frac{-ie\mathcal{E}_{ty}}{\hbar}\mathrm{e}^{(i\omega_h-\lambda_1)\tau}\, \langle hr|\, y\, |cq\rangle\, [a^\dagger_{hr}a_{hr}(-\tau) - a^\dagger_{cq}a_{cq}(-\tau)] \\ a^\dagger_{cr}a_{lq}(0^-) \quad &= \quad \frac{-ie\mathcal{E}_{ty}}{\hbar}\mathrm{e}^{(i\omega_l-\lambda_1)\tau}\, \langle lq|\, y\, |cr\rangle\, [a^\dagger_{lq}a_{lq}(-\tau) - a^\dagger_{cr}a_{cr}(-\tau)]. \end{aligned} \quad (6.22)$$

Supposing that the valence bands are initially completely filled, we further omit the last factor in these expressions. Inserting these into the operators at time (0^+) and substituting these again in the the expression (6.17) we obtain for that contribution to the total y-polarization (6.13) the result, intentionally written in this form,

$$Y_{h\vec{k}}(t>0) \quad = \quad \frac{ie^3\mathcal{E}_p^2\mathcal{E}_{ty}}{\hbar^3}\mathrm{e}^{-\lambda_2 t}\mathrm{e}^{-\lambda_1\tau}\mathrm{e}^{-i\omega_h t}\mathcal{U} \tag{6.23}$$

with

$$\begin{aligned}
\mathcal{U} \quad = \quad \sum_{ijrq} \langle hi,\vec{k}|\, y\, |cj,\vec{k}\rangle \Big[& \langle cj,\vec{k}|\, x\, |hr,\vec{k}\rangle \langle hr,\vec{k}|\, x\, |cq,\vec{k}\rangle \langle cq,\vec{k}|\, y\, |hi,\vec{k}\rangle \,\mathrm{e}^{-i\omega_h\tau} \\
& + \langle cj,\vec{k}|\, x\, |hr,\vec{k}\rangle \langle hr,\vec{k}|\, y\, |cq,\vec{k}\rangle \langle cq,\vec{k}|\, x\, |hi,\vec{k}\rangle \,\mathrm{e}^{i\omega_h\tau} \\
& + \langle cr,\vec{k}|\, x\, |hi,\vec{k}\rangle \langle hq,\vec{k}|\, x\, |cr,\vec{k}\rangle \langle cj,\vec{k}|\, y\, |hq,\vec{k}\rangle \,\mathrm{e}^{-i\omega_h\tau} \\
& + \langle cr,\vec{k}|\, x\, |hi,\vec{k}\rangle \langle hq,\vec{k}|\, y\, |cr,\vec{k}\rangle \langle cj,\vec{k}|\, x\, |hq,\vec{k}\rangle \,\mathrm{e}^{i\omega_h\tau} \\
& + \langle cr,\vec{k}|\, x\, |hi,\vec{k}\rangle \langle lq,\vec{k}|\, x\, |cr,\vec{k}\rangle \langle cj,\vec{k}|\, y\, |lq,\vec{k}\rangle \,\mathrm{e}^{-i\omega_l\tau} \\
& + \langle cr,\vec{k}|\, x\, |hi,\vec{k}\rangle \langle lq,\vec{k}|\, y\, |cr,\vec{k}\rangle \langle cj,\vec{k}|\, x\, |lq,\vec{k}\rangle \,\mathrm{e}^{i\omega_l\tau} \\
& + \langle cj,\vec{k}|\, x\, |lr,\vec{k}\rangle \langle lr,\vec{k}|\, x\, |cq,\vec{k}\rangle \langle cq,\vec{k}|\, y\, |hi,\vec{k}\rangle \,\mathrm{e}^{-i\omega_h\tau} \\
& + \langle cj,\vec{k}|\, x\, |lr,\vec{k}\rangle \langle lr,\vec{k}|\, y\, |cq,\vec{k}\rangle \langle cq,\vec{k}|\, x\, |hi,\vec{k}\rangle \,\mathrm{e}^{i\omega_l\tau}].\Big].
\end{aligned}$$

This form is quite instructive, because the various terms can be interpreted as a "round trip" of the electron between the bands. Two factors are matrix elements that describe the promotion of an electron from a heavy-hole or light-hole band, by x-polarized or y-polarized light, to the conduction band; the other two factors describe the return to these bands. In some of the terms the steps "upward", from valence band to conduction band, are both done by the x operator and the steps "downward", from conduction band to valence band, by the y operator, or just the opposite. In other terms the x operator and the y operator each occur in one step upward and one step downward. Since the matrix element that describes the step in one direction is just the complex conjugate of the matrix element that describes the opposite step, for instance $\langle hi,\vec{k}|\, y\, |cj,\vec{k}\rangle = \langle cj,\vec{k}|\, y\, |hi,\vec{k}\rangle^*$, and because the y occurs always with an imaginary factor in the basis functions, in contrast to x, we see that loosely speaking the electron pics up twice a factor i or twice a factor $-i$, that is a minus sign, in some of its "round trips", but a factor i times a factor $-i$ in other terms. This does not play a role in the case that pump and probe are both x-polarized, where we have only steps made with the x operator:

$$X_{h\vec{k}}(t>0) = \frac{ie^3\mathcal{E}_p^2\mathcal{E}_{tx}}{\hbar^3} e^{-\lambda_2 t}\mathrm{e}^{-\lambda_1\tau}e^{-i\omega_h t}\mathcal{V} \tag{6.24}$$

with

$$\begin{aligned}
\mathcal{V} = \sum_{ijrq} \langle hi,\vec{k}|\, x\,|cj,\vec{k}\rangle \Big[&\langle cj,\vec{k}|\, x\,|hr,\vec{k}\rangle \langle hr,\vec{k}|\, x\,|cq,\vec{k}\rangle \langle cq,\vec{k}|\, x\,|hi,\vec{k}\rangle\, \mathrm{e}^{-i\omega_h\tau} \\
&+ \langle cj,\vec{k}|\, x\,|hr,\vec{k}\rangle \langle hr,\vec{k}|\, x\,|cq,\vec{k}\rangle \langle cq,\vec{k}|\, x\,|hi,\vec{k}\rangle\, \mathrm{e}^{i\omega_h\tau} \\
&+ \langle cr,\vec{k}|\, x\,|hi,\vec{k}\rangle \langle hq,\vec{k}|\, x\,|cr,\vec{k}\rangle \langle cj,\vec{k}|\, x\,|hq,\vec{k}\rangle\, \mathrm{e}^{-i\omega_h\tau} \\
&+ \langle cr,\vec{k}|\, x\,|hi,\vec{k}\rangle \langle hq,\vec{k}|\, x\,|cr,\vec{k}\rangle \langle cj,\vec{k}|\, x\,|hq,\vec{k}\rangle\, \mathrm{e}^{i\omega_h\tau} \\
&+ \langle cr,\vec{k}|\, x\,|hi,\vec{k}\rangle \langle lq,\vec{k}|\, x\,|cr,\vec{k}\rangle \langle cj,\vec{k}|\, x\,|lq,\vec{k}\rangle\, \mathrm{e}^{-i\omega_l\tau} \\
&+ \langle cr,\vec{k}|\, x\,|hi,\vec{k}\rangle \langle lq,\vec{k}|\, x\,|cr,\vec{k}\rangle \langle cj,\vec{k}|\, x\,|lq,\vec{k}\rangle\, \mathrm{e}^{i\omega_l\tau} \\
&+ \langle cj,\vec{k}|\, x\,|lr,\vec{k}\rangle \langle lr,\vec{k}|\, x\,|cq,\vec{k}\rangle \langle cq,\vec{k}|\, x\,|hi,\vec{k}\rangle\, \mathrm{e}^{-i\omega_h\tau} \\
&+ \langle cj,\vec{k}|\, x\,|lr,\vec{k}\rangle \langle lr,\vec{k}|\, x\,|cq,\vec{k}\rangle \langle cq,\vec{k}|\, x\,|hi,\vec{k}\rangle\, \mathrm{e}^{i\omega_l\tau} \Big].
\end{aligned}$$

The consequence is that for parallel polarization all terms add up in a constructive manner, while they partially cancel for the case of orthogonal polarization. The FWM signal is thereby stronger in case of parallel polarization of pump and probe. Somewhat closer inspection is needed to see that for delay time $\tau = 0$ the FWM signal has a maximum in case of parallel polarization of both pulses, while there is just a minimum for the cross polarized case.

Adding the contribution $X_{l\vec{k}}(t>0)$, which is just (6.24) with h and l interchanged, we obtain

$$\begin{aligned}
P_{\parallel}^{(3)}(t) = \; &\Theta(t) i \frac{e^3\mathcal{E}_p^2\mathcal{E}_{tx}}{\hbar^3}\mathrm{e}^{-\lambda_1\tau}\mathrm{e}^{-\lambda_2 t} \\
&\Big[\mathrm{e}^{-i\omega_h t}\{2A\mathrm{e}^{-i\omega_h\tau} + (2A+B)\mathrm{e}^{i\omega_h\tau} + 2B\mathrm{e}^{-i\omega_l\tau} + B\mathrm{e}^{i\omega_l\tau}\} \\
&+\mathrm{e}^{-i\omega_l t}\{2B\mathrm{e}^{-i\omega_h\tau} + B\mathrm{e}^{i\omega_h\tau} + 2C\mathrm{e}^{-i\omega_l\tau} + (2C+B)\mathrm{e}^{i\omega_l\tau}\}\Big]
\end{aligned}$$

with

$$\begin{aligned}
A &= \sum_{\vec{k}}\sum_{ijrq} \langle hi,\vec{k}|\, x\,|cj,\vec{k}\rangle \langle cj,\vec{k}|\, x\,|hr,\vec{k}\rangle \langle hr,\vec{k}|\, x\,|cq,\vec{k}\rangle \langle cq,\vec{k}|\, x\,|hi,\vec{k}\rangle \\
&= \sum_{\vec{k}} \frac{1}{2N^2}(|b|^2 + |R_h\frac{1}{\sqrt{3}} + c|^2)^2
\end{aligned}$$

$$\begin{aligned}
B &= \sum_{\vec{k}}\sum_{ijrq} \langle hi,\vec{k}|\, x\, |cj,\vec{k}\rangle\, \langle cj,\vec{k}|\, x\, |lr,\vec{k}\rangle\, \langle lr,\vec{k}|\, x\, |cq,\vec{k}\rangle\, \langle cq,\vec{k}|\, x\, |hi,\vec{k}\rangle \\
&= \sum_{\vec{k}} \frac{1}{6N^2}(|b|^2 + |R_h\frac{1}{\sqrt{3}} + c|^2)(|b|^2 + |R_l\sqrt{3} - c|^2)
\end{aligned}$$

$$\begin{aligned}
C &= \sum_{\vec{k}}\sum_{ijrq} \langle li,\vec{k}|\, x\, |cj,\vec{k}\rangle\, \langle cj,\vec{k}|\, x\, |lr,\vec{k}\rangle\, \langle lr,\vec{k}|\, x\, |cq,\vec{k}\rangle\, \langle cq,\vec{k}|\, x\, |li,\vec{k}\rangle \\
&= \sum_{\vec{k}} \frac{1}{18N^2}(|b|^2 + |R_l\sqrt{3} - c|^2)^2.
\end{aligned}$$

The summation over $\vec{k}$ can be split into a summation over $|k|$ and an integration over the angles of $\vec{k}$. The latter can easily be done analytically, with $N = 3k^2\sin^2\theta$,we obtain:

$$\begin{aligned}
\int d\Omega_k \frac{1}{2N^2}(|b|^2 + |R_h\frac{1}{\sqrt{3}} + c|^2)^2 &= \frac{48}{45}\pi \\
\int d\Omega_k \frac{1}{6N^2}(|b|^2 + |R_h\frac{1}{\sqrt{3}} + c|^2)(|b|^2 + |R_l\sqrt{3} - c|^2) &= \frac{32}{45}\pi \\
\int d\Omega_k \frac{1}{18N^2}(|b|^2 + |R_l\sqrt{3} - c|^2)^2 &= \frac{48}{45}\pi.
\end{aligned}$$

One should keep in mind that the frequencies ω_h and ω_l depend in principle on $|k|$, that is on the part of the band that is involved in the excitation. In practice this is always a certain range of frequencies; even in the case that bound excitons are sticking out, as for quantum wells at low temperature. Therefore the fast oscillations as a function of the delay time τ that are proportional to $\mathrm{e}^{\pm 2i\omega_h\tau}$, $\mathrm{e}^{\pm 2i\omega_l\tau}$ and $\mathrm{e}^{\pm i(\omega_h+\omega_l)\tau}$ will be averaged out due to the summation over $|k|$ in the final result (6.6) and only the much more slowly varying exponential function of the beat frequency $\omega_l - \omega_h$ survives. The time integrals over the exponential functions are elementary and we obtain for the detected intensity of the FWM signal in case of parallel polarization:

$$\begin{aligned}
I_{\parallel} &= \frac{\mathcal{E}_p^4\mathcal{E}_{tx}^2}{\hbar^6}\left(\frac{32\pi}{45}\right)^2 \mathrm{e}^{-2\lambda_1\tau} F \frac{1}{2\lambda_2}\Bigg[60 + 40\cos(\omega_h - \omega_l)\tau + \frac{4\lambda_2^2}{4\lambda_2^2 + (\omega_h - \omega_l)^2} \\
&\quad \{40 + 60\cos(\omega_h - \omega_l)\tau\} + \frac{2\lambda_2(\omega_h - \omega_l)}{4\lambda_2^2 + (\omega_h - \omega_l)^2} 20\sin(\omega_h - \omega_l)\tau\Bigg], \qquad (6.25)
\end{aligned}$$

in which the factor F accounts for the summation over $|k|$.

The case of orthogonal polarization can be treated in the same way:

$$\begin{aligned} P_{\perp}^{(3)}(t) &= \Theta(t) i \frac{\mathcal{E}_p^2 \mathcal{E}_{ty}}{\hbar^3} e^{-\lambda_1 \tau} e^{-\lambda_2 t} \\ &\quad \Big[\mathrm{e}^{-i\omega_h t} (\alpha_1 \mathrm{e}^{-i\omega_h \tau} + \alpha_2 \mathrm{e}^{i\omega_h \tau} + \alpha_3 \mathrm{e}^{-i\omega_l \tau} + \alpha_4 e^{i\omega_l \tau}) \\ &\quad + \mathrm{e}^{-i\omega_l t} (\beta_1 \mathrm{e}^{-i\omega_h \tau} + \beta_2 \mathrm{e}^{i\omega_h \tau} + \beta_3 \mathrm{e}^{-i\omega_l \tau} + \beta_4 \mathrm{e}^{i\omega_l \tau}) \Big], \end{aligned}$$

with

$$\begin{aligned} \alpha_1 &= \sum_{\vec{k}} \sum_{ijrq} \langle hi, \vec{k}| \, y \, |cj, \vec{k}\rangle \langle cj, \vec{k}| \, x \, |hr, \vec{k}\rangle \langle hr, \vec{k}| \, x \, |cq, \vec{k}\rangle \langle cq, \vec{k}| \, y \, |hi, \vec{k}\rangle \\ &+ \sum_{\vec{k}} \sum_{ijrq} \langle hi, \vec{k}| \, y \, |cj, \vec{k}\rangle \langle cj, \vec{k}| \, y \, |hr, \vec{k}\rangle \langle hr, \vec{k}| \, x \, |cq, \vec{k}\rangle \langle cq, \vec{k}| \, x \, |hi, \vec{k}\rangle \\ &+ \sum_{\vec{k}} \sum_{ijrq} \langle hi, \vec{k}| \, y \, |cj, \vec{k}\rangle \langle cj, \vec{k}| \, x \, |lr, \vec{k}\rangle \langle lr, \vec{k}| \, x \, |cq, \vec{k}\rangle \langle cq, \vec{k}| \, y \, |hi, \vec{k}\rangle , \end{aligned}$$

$$\begin{aligned} \alpha_2 &= 2 \sum_{\vec{k}} \sum_{ijrq} \langle hi, \vec{k}| \, y \, |cj, \vec{k}\rangle \langle cj, \vec{k}| \, x \, |hr, \vec{k}\rangle \langle hr, \vec{k}| \, y \, |cq, \vec{k}\rangle \langle cq, \vec{k}| \, x \, |hi, \vec{k}\rangle , \\ \alpha_3 &= \sum_{\vec{k}} \sum_{ijrq} \langle hi, \vec{k}| \, y \, |cj, \vec{k}\rangle \langle cj, \vec{k}| \, y \, |lr, \vec{k}\rangle \langle lr, \vec{k}| \, x \, |cq, \vec{k}\rangle \langle cq, \vec{k}| \, x \, |hi, \vec{k}\rangle , \\ \alpha_4 &= 2 \sum_{\vec{k}} \sum_{ijrq} \langle hi, \vec{k}| \, y \, |cj, \vec{k}\rangle \langle cj, \vec{k}| \, x \, |lr, \vec{k}\rangle \langle lr, \vec{k}| \, y \, |cq, \vec{k}\rangle \langle cq, \vec{k}| \, x \, |hi, \vec{k}\rangle , \end{aligned}$$

$$\beta_1 = \sum_{\vec{k}} \sum_{ijrq} \langle li, \vec{k}| \, y \, |cj, \vec{k}\rangle \langle cj, \vec{k}| \, y \, |hr, \vec{k}\rangle \langle hr, \vec{k}| \, x \, |cq, \vec{k}\rangle \langle cq, \vec{k}| \, x \, |li, \vec{k}\rangle ,$$

$$\beta_2 = 2 \sum_{\vec{k}} \sum_{ijrq} \langle li, \vec{k}| \, y \, |cj, \vec{k}\rangle \langle cj, \vec{k}| \, x \, |hr, \vec{k}\rangle \langle hr, \vec{k}| \, y \, |cq, \vec{k}\rangle \langle cq, \vec{k}| \, x \, |li, \vec{k}\rangle ,$$

$$\begin{aligned} \beta_3 &= \sum_{\vec{k}} \sum_{ijrq} \langle li, \vec{k}| \, y \, |cj, \vec{k}\rangle \langle cj, \vec{k}| \, x \, |lr, \vec{k}\rangle \langle lr, \vec{k}| \, x \, |cq, \vec{k}\rangle \langle cq, \vec{k}| \, y \, |li, \vec{k}\rangle , \\ &+ \sum_{\vec{k}} \sum_{ijrq} \langle li, \vec{k}| \, y \, |cj, \vec{k}\rangle \langle cj, \vec{k}| \, y \, |lr, \vec{k}\rangle \langle lr, \vec{k}| \, x \, |cq, \vec{k}\rangle \langle cq, \vec{k}| \, x \, |li, \vec{k}\rangle , \\ &+ \sum_{\vec{k}} \sum_{ijrq} \langle li, \vec{k}| \, y \, |cj, \vec{k}\rangle \langle cj, \vec{k}| \, x \, |hr, \vec{k}\rangle \langle hr, \vec{k}| \, x \, |cq, \vec{k}\rangle \langle cq, \vec{k}| \, y \, |li, \vec{k}\rangle , \end{aligned}$$

$$\beta_4 = 2\sum_{\vec{k}}\sum_{ijrq} \langle li,\vec{k}|\, y\,|cj,\vec{k}\rangle \,\langle cj,\vec{k}|\, x\,|lr,\vec{k}\rangle \,\langle lr,\vec{k}|\, y\,|cq,\vec{k}\rangle \,\langle cq,\vec{k}|\, x\,|li,\vec{k}\rangle \,.$$

Expressed in terms of the coefficients b, c, R_i of the band wave functions we obtain

$$\begin{aligned}\alpha_1 \;=\; & \sum_{\vec{k}} \frac{1}{2N^2}\Bigg[(|b|^2+|R_h\frac{1}{\sqrt{3}}+c|^2)(|b|^2+|R_h\frac{1}{\sqrt{3}}-c|^2)+(|b|^2+|R_h\frac{1}{\sqrt{3}}-c|^2)\\ & (|b|^2+|R_h\frac{1}{\sqrt{3}}+c|^2)+\frac{1}{3}(|b|^2+|R_h\frac{1}{\sqrt{3}}-c|^2)(|b|^2+|R_l\sqrt{3}-c|^2)\Bigg],\end{aligned}$$

$$\alpha_2 = -\sum_{\vec{k}} \frac{1}{N^2}\Bigg[(|b|^2+|c|^2-\frac{1}{3}R_h^2)^2+\frac{1}{3}R_h^2(c-c^*)^2+\frac{4}{3}|b|^2R_h^2\Bigg],$$

$$\alpha_3 = -\sum_{\vec{k}} \frac{1}{6N^2}\Bigg[(|b|^2+|c|^2-3R_l^2)(|b|^2+|c|^2-\frac{1}{3}R_h^2)-R_hR_l(c-c^*)^2+4|b|^2R_hR_l\Bigg],$$

$$\alpha_4 = \sum_{\vec{k}} \frac{1}{6N^2}\Bigg[(|b|^2+|c|^2-3R_l^2)(|b|^2+|c|^2-\frac{1}{3}R_h^2)+R_hR_l(c-c^*)^2+4|b|^2R_hR_l\Bigg],$$

$$\beta_1 = {\alpha_3}^*,$$
$$\beta_2 = \alpha_4,$$

$$\begin{aligned}\beta_3 \;=\; & \sum_{\vec{k}} \frac{1}{6N^2}\Bigg[\frac{1}{3}(|b|^2+|R_l\sqrt{3}-c|^2)(|b|^2+|R_l\sqrt{3}+c|^2)+\frac{1}{3}(|b|^2+|R_l\sqrt{3}+c|^2)\\ & (|b|^2+|R_l\sqrt{3}-c|^2)+(|b|^2+|R_h\frac{1}{\sqrt{3}}+c|^2)(|b|^2+|R_l\sqrt{3}+c|^2)\Bigg],\end{aligned}$$

$$\beta_4 = -\sum_{\vec{k}} \frac{1}{18N^2}\Bigg[(|b|^2+|c|^2-3R_l^2)^2+3R_l^2(c-c^*)^2+12|b|^2R_l^2\Bigg].$$

The integrals over the angles of $\vec{k}$ now yield the factors:

$$\begin{aligned}&\text{for}\quad \alpha_1 \;:\; \frac{36}{45}\pi+\frac{36}{45}\pi+\frac{44}{45}\pi; \qquad \text{for}\quad \alpha_2 : -\frac{48}{45}\pi;\\ &\text{for}\quad \alpha_3 \;:\; -\frac{16}{45}\pi; \qquad \text{for}\quad \alpha_4 : \frac{8}{45}\pi\\ &\text{for}\quad \beta_1 \;:\; -\frac{16}{45}\pi; \qquad \text{for}\quad \beta_2 : \frac{8}{45}\pi\\ &\text{for}\quad \beta_3 \;:\; \frac{36}{45}\pi+\frac{36}{45}\pi+\frac{44}{45}\pi; \qquad \text{for}\quad \beta_4 : -\frac{48}{45}\pi\end{aligned}$$

and the final result becomes

$$\begin{aligned}I_{\perp} \;=\; & \frac{\mathcal{E}_p^4\mathcal{E}_{ty}^2}{\hbar^6}\left(\frac{32\pi}{45}\right)^2 \mathrm{e}^{-2\lambda_1\tau}\frac{F}{2\lambda_2}\left[31.4-8.75\cos(\omega_h-\omega_l)\tau+\frac{4\lambda_2^2}{4\lambda_2^2+(\omega_h-\omega_l)^2}\right. \\ & \left.(-8.75+31.4\cos(\omega_h-\omega_l)\tau)-\frac{2\lambda_2(\omega_h-\omega_l)}{4\lambda_2^2+(\omega_h-\omega_l)^2}10.66\mathrm{sin}(\omega_h-\omega_l)\tau\right].\end{aligned} \tag{6.26}$$

The value of λ_1 can be clearly read off from the experimental decay of the FWM signal as a function of the delay time τ. If one assumes $\lambda_2 \approx \lambda_1$ and takes also the beat frequency from the data, one verifies that the first term in the expressions (6.25) and (6.26) dominates. These then show three features that are also observed in experiments. Firstly the FWM signal is stronger in the case of parallel polarization then in case of cross polarization, though we find only a factor 2 while the experiment of Bennhardt *et al.* [117] gave approximately a factor 4 difference in strength. Earlier theoretical expressions [116] gave equal strength in both cases, obviously due to omission of some contributions. Secondly, we do indeed find a beat behavior as a function of the delay time which has a maximum at delay time zero for the case of parallel polarization and just the opposite for orthogonal polarization. And thirdly, we find that the beats are more pronounced for parallel polarization than for orthogonal polarization. We therefore conclude that the main features seen in experiments are to a considerable extent described without invoking other mechanisms than just the coupling of the carrier (polarization) dynamics with the light fields.

The calculation of Fig.(6.1) is performed adopting an ad-hoc relaxation time $\lambda = 4$ps, to account for the diffusion into the non-collective space by Coulomb collisions. For the wavelength of the optical transition $|g\rangle \rightarrow |h\rangle$ we adopted 810 nm and a difference of the energy between $|h\rangle$ and $|l\rangle$ states 4.2 meV.

6.3 Conclusion

In conclusion, we have re-examined the description of quantum beats in four-wave mixing in bulk semiconductor and found that the main features may be described by including the coherent coupling between the light field and the charge carriers, without introducing the exciton-exciton coupling by the Coulomb interaction between the charge carriers, in contrast to earlier theoretical work. The equations of motion for the polarization involve coupling terms between heavy-hole and light-hole states that usually do not appear in

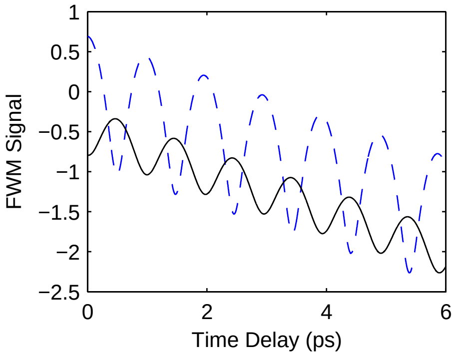

Figure 6.1: The FWM signal ($Ln(I)$) quantum beats as a function of pump-probe time delay. The dash-dotted and solid lines for parallel and orthogonal polarized pump-probe configurations.

the standard semiconductor Bloch equations (SBE), but are similar to those in the treatment of optical Bloch equations (OBE). We presented explicit equations for the dipole polarization in terms of products of single-particle matrix elements which expose in a transparent way why the FWM signal has different strength for co-polarized or cross-polarized pulses. Also the different phases of the beats as a function of the delay time are reproduced. So we argue that the observed beat phenomena are largely due to the coherent light-carrier interaction process.

Bibliography

[1] G. P. Agrawal, Fiber optical communication systems. New York: Wiley, 2002.

[2] H. Haug and S. W. Koch, Quantum Theory of the Optical and Electronic Properties of Semiconductors, 3 ed. Singapore: World Scientific, 1994.

[3] D. Cotter, R. J. Manning, K. J. Blow, A. D. Ellis, A. E. Kelly, D. Nesset, I. D. Phillips, A. J. Poustie, and D. C. Rogers, "Nonlinear optics for high-speed digital information processing," Science, vol. 286, p. 1523, 1999.

[4] R. W. Boyd, Nonlinear Optics. New York: Academic Press, 1992.

[5] A. Yariv, Optical Electronics in Modern Communications, 5 ed. London: Oxford University Press, 1997.

[6] D. Suter, The Physics of Laser-Atom Interactions. USA: Cambridge University Press, 1997.

[7] W. W. Chow, S. W. Koch, and M. S. Sargent III, Semiconductor-laser physics. Berlin: Springer, 1994.

[8] S. M. Jensen, "The non-linear coherent coupler " IEEE J. Quantum. Electron., vol. 18, pp. 1580-1583, 1982.

[9] Y. Wang and W. Wang, "Pulse-width selective all-optical switching," Appl. Phys. Lett., vol. 88, p. 181110, 2006.

[10] A. T. Pham and L. N. Binh, "All-optical modulation and switching using a nonlinear-optical directional coupler," J. Opt. Soc. Am. B, vol. 8, p. 1914, 1991.

[11] K.-i. Kitayama and S. Wang, "Optical pulse compression by nonlinear coupling," Appl. Phys. Lett., vol. 43, pp. 17-19, 1983.

[12] Y. Chen, A. W. Snyder, and D. J. Mitchell, "Ideal optical switching by nonlinear multiple (parasitic) core couplers," Electron. Lett., vol. 26, pp. 77-78, 1990.

[13] Y. Wang, "Nonlinear optical limiter and digital optical switch by cascaded nonlinear couplers: analysis," IEEE/OSA J. Lightwave Tech., vol. 12, p. 292, 1999.

[14] Y. Wang and J. Liu, "All-fiber logical devices based on the nonlinear directional coupler," IEEE J. Photon. Tech. Lett., vol. 11, pp. 72-74, 1999.

[15] C. Thirstrup, "Optical bistability in a nonlinear directional coupler," IEEE J. Quantum. Electron., vol. 31, pp. 2101-2106, 1995.

[16] S. R. Friberg, Y. Silberberg, M. K. Oliver, M. J. Andrejco, M. A. Saifi, and P. W. Smith, "Ultrafast all-optical switching in a dual-core fiber nonlinear coupler " Appl. Phys. Lett., vol. 51, p. 1135, 1987.

[17] A. Villeneuve, C. C. Yang, P. G. J. Wigley, G. I. Stegeman, J. S. Aitchison, and C. N. Ironside, "Ultrafast all-optical switching in semiconductor nonlinear directional couplers at half the band gap," Appl. Phys. Lett., vol. 61, p. 147, 1992.

[18] K. Al-hemyari, A. Villeneuve, J. U. Kang, J. S. Aitchison, C. N. Ironside, and G. I. Stegeman, "Ultrafast all-optical switching in GaAlAs directional couplers at 1.55 m without multiphoton absorption," Appl. Phys. Lett., vol. 63, p. 3562, 1993.

[19] J. S. Aitchison, A. Villeneuve, and G. I. Stegeman, "All-optical switching in two cascaded nonlinear directional couplers," Opt. Lett., vol. 20, p. 698, 1994.

[20] S. R. Friberg, A. M. Weiner, Y. Silberberg, B. G. Sfez, and P. S. Smith, "Femtosecond switching in a dual-core-fiber nonlinear coupler," Opt. Lett., vol. 13, 1988.

[21] Y. Wang and W. Wang, "Nonlinear optical pulse coupling dynamics " IEEE/OSA J. Lightwave Tech., vol. 24, pp. 2458-2464, 2006.

[22] Y. Wang and W. Wang, "Study of ultrafast pulse coupling dynamics considering retarded nonlinear response and self-steepening effects," IEEE/OSA J. Lightwave Tech., vol. 24, pp. 1041-1047, 2006.

[23] Y. Wang and W. Wang, "Study of ultra-fast optical pulse propagation in a nonlinear directional coupler," Appl. Phys. B, vol. 79, p. 51, 2004.

[24] Y. Wang and C. K. Lee, "Effects of the linear mismatch and nonlinear asymmetry in the nonlinear coupler " Appl. Phys. B, vol. 72, p. 417, 2001.

[25] Y. Wang and W. Wang, "A Simple and Efficient Numerical Method for Nonlinear Pulse Propagation in N-Core Optical Couplers," IEEE Photon. Tech. Lett., vol. 16, 2004.

[26] S. Trillo, S. Wabnitz, E. M. Wright, and G. I. Stegeman, "Soliton switching in fiber nonlinear directional couplers," Opt. Lett., vol. 13, p. 672, 1988.

[27] P. L. Chu, B. A. Malomed, and G.-D. Peng, "Soliton switching and propagation in nonlinear fiber couplers: analytical results," J. Opt. Soc. Am. B, vol. 10, p. 1379, 1993.

[28] Y. S. Kivshar, "Switching dynamics of solitons in fiber directional couplers," Opt. Lett., vol. 18, p. 7, 1993.

[29] C. R. Paiva, A. L. Topa, and A. M. Barbosa, "Influence of intermodal dispersion on the switching of solitons at different wavelengths in twin-core fiber couplers " J. Opt. Soc. Am. B, vol. 16, p. 1636, 1999.

[30] K. S. Chiang, "Intermodal dispersion in two-core optical fibers," Opt. Lett., vol. 20, p. 997, 1995.

[31] P. M. Ramos and C. R. Paiva, "All-optical pulse switching in twin-core fiber couplers with intermodal dispersion," IEEE J. Quantum. Electron., vol. 35, pp. 983-989, 1999.

[32] D. S. Citrin, "Self-Pulse-Shaping Coherent Control of Excitons in a Semiconductor Microcavity," Phys. Rev. Lett., vol. 77, p. 4596, 1996.

[33] G. S. He and S. H. Liu, Physics of Nonlinear Optics. Singapore: World Scientific, 1999.

[34] J. P. Heritage, A. M. Weiner, and R. N. Thurston, "Picosecond pulse shaping by spectral phase and amplitude manipulation," Opt. Lett., vol. 10, p. 609, 1985.

[35] P. Petropoulos, M. Ibsen, A. D. Ellis, and D. J. Richardson, "Rectangular pulse generation based on pulse reshaping using a superstructured fiber Bragg grating," IEEE/OSA J. Lightwave Tech., vol. 19, pp. 746-752, 2001.

[36] A. M. Weiner, Y. Silberberg, H. Fouckhardt, D. E. Leaird, M. A. Saifi, M. J. Andrejco, and P. W. Smith, "Use of femtosecond square pulses to avoid pulse breakup in all-optical switching," IEEE J. Quantum. Electron., vol. 25, p. 2648, 1989.

[37] R. Hoffe and J. Chrostowski, "Optical pulse compression and breaking in nonlinear fibre couplers " Opt. Commun., vol. 57, p. 34, 1986.

[38] K. Yasumoto, H. Maeda, and N. Maekawa, "Coupled-mode analysis of an asymmetric nonlinear directional coupler," IEEE/OSA J. Lightwave Tech., vol. 14, pp. 628-633, 1996.

[39] J. Atai and Y. Chen, "Nonlinear mismatches between two cores of saturable nonlinear couplers," IEEE J. Quantum. Electron., vol. 29, pp. 242-249, 1993.

[40] Y. Chen, "Twin core nonlinear couplers with saturable nonlinearity," Electron. Lett., vol. 26, p. 1374, 1990.

[41] S.-H. Cho, H. Kumagai, and K. Midorikawa, "Fabrication of multi-core structures in an optical fiber using plasma self-channeling," vol. 11, pp. 1780-1786, 2003.

[42] D.J. Hagan, T. Xia, A.A. Said, E.W.Van Stryland; SPIE 2229, 179, 1994

[43] G.R. Olbright and N. Peyghambarian; Appl. Phys. Lett. 48, 1184, 1986

[44] A.T. Pham and L.N. Binh; Int. J. Optoelectron. 5, 367, 1990

[45] A.W. Snyder and D. Love, Optical Waveguide Theory, London, Chapman & Hall, 1983

[46] S. Trillo and S. Wabnitz, "Nonlinear nonreciprocity in a coherent mismatched directional coupler," Appl. Phys. Lett., vol. 49, p. 752, 1986.

[47] W. Wang, Y. Wang, K. Allaart, and D. Lenstra, "Ultrashort rectangular optical pulse generation by nonlinear directional couplers," Opt. Commun., vol. 253, p. 164, 2005.

[48] C.-C. Yang and A. J. S. Wang, "Asymmetric nonlinear coupling and its applications to logic functions," IEEE J. Quantum. Electron., vol. 28, pp. 479-487, 1992.

[49] A. R. Chraplyvy, A. H. Gnauck, R. W. Tkach, J. L. Zyskind, J. W. Sulhoff, A. J. Lucero, Y. Sun, R. M. Jopson, F. Forghieri, R. M. Derosier, C. Wolf, and A. R. McCormick, "1-Tb/s transmission experiment," IEEE Photon. Tech. Lett., vol. 8, p. 1264, 1996.

[50] S. Kawanishi, H. Takara, K. Uchiyama, I. Shake, O. Kamatani, and H. Takahashi, "1.4 Tbit/s (200 Gbit/s × 7 ch) 50 km optical transmission experiment," Electron. Lett., vol. 33, p. 1716, 1997.

[51] M. Nakazawa, E. Yoshida, T. Yamamoto, E. Yamada, and A. Sahara, "TDM single channel 640 Gbit/s transmission experiment over 60 km using 400 fs pulse train and walk-off free, dispersion flattened nonlinear optical loop mirror," Electron. Lett., vol. 34, p. 907, 1998.

[52] K. E. Stubkjaer, "Semiconductor optical amplifier-based all-optical gates for high-speed optical processing," IEEE J. Select. Topics Quantum Electron., vol. 6, p. 1428, 2000.

[53] J. C. Slater and G. F. Koster, "Simplified LCAO method of the periodic potential problem," Phys. Rev., vol. 94, p. 1498, 1954.

[54] M. F. Li, Modern semiconductor quantum physics. Singapore: World Scienfic, 1994.

[55] D. Birkedal, "Electronic structure and optical properties of semiconductor heterostructures," in COM. vol. PhD Copenhagen: Denmark Tech. Univ., 1991.

[56] M. Persson, "Theoretical study of electronic structure and optical properties of semiconductor nanostructures." vol. PhD: Lund Univ., 2004.

[57] E. O. Kane, Handbook on semiconductors vol. 1. Amsterdam: North Holland, 1982.

[58] W. A. Harrison, Solid state theory. New York: McGraw-Hill, 1970.

[59] P. O. Loewdin, " A note on the quantum-mechanical perturbation theory," J. Chem. Phys., vol. 19, p. 1386, 1951.

[60] J. Luttinger and W. Kohn, "Motion of electrons and holes in perturbed periodic fields," Phys. Rev., vol. 97, p. 869, 1955.

[61] I. Suemune, L. A. Coldren, M. Yamanishi, and Y. Kan, "Extremely wide modulation bandwidth in a low threshold current strained quantum well laser," Appl. Phys. Lett., vol. 53, p. 1378, 1988.

[62] A. Ghiti, E. P. O'Reilly, and A. R. Adams, "Improved dynamics and linewidth enhancement factor in strained-layer lasers " Electron. Lett., vol. 25, p. 821, 1989.

[63] K. Y. Lau, S. Xin, W. I. Wang, N. Bar-Chaim, and M. Mittelstein, "Enhancement of modulation bandwidth in InGaAs strained-layer single quantum well lasers," Appl. Phys. Lett., vol. 55, p. 1973, 1989.

[64] G. Jones and E. P. O'Reilly, "Improved performance of long-wavelength strained bulk-like semiconductor lasers," IEEE J. Quantum. Electron., vol. 29, pp. 1344-1354, 1993.

[65] E. P. O'Reilly and A. R. Adams, "Band-structure engineering in strained semiconductor lasers," IEEE J. Quantum. Electron., vol. 30, p. 366, 1994.

[66] G. L. Bir and G. E. Pikus, Symmetry and strain-induced effects in semiconductor. New York: Wiley, 1972.

[67] O. Madelung, Landolt-Boernstein New series, Group III, 17. Berlin: Springer, 1982.

[68] T. Saitoh and T. Mukai, "1.5 μm GaInAsP traveling-wave semiconductor laser amplifier" IEEE J. Quantum. Electron., vol. 23, p. 1010, 1978.

[69] J. Mork, M. L. Nielsen, and T. W. Berg, "The dynamics of semiconductor optical amplifiers: modeling and applications," Optics & Photonics News, Jul. , p. 43, 2003.

[70] M. J. Connelly, Semiconductor optical amplifiers: Kluwer Academic Publishers, Amsterdam, 2002.

[71] R. H. Yan, S. W. Corzine, L. A. C. and, and I. Suemune, "Corrections to the expression for gain in GaAs," IEEE J. Quantum. Electron., vol. 26, p. 213, 1990.

[72] J. Singh, Semiconductor Optoelectronics: Physics and Technology. New York: McGraw-Hill, 1995.

[73] M. Asada, A. R. Admas, K. E. Stubkjaer, Y. Suematsu, Y. Itaha, and S. Arai, "The temperature dependence of the threshold current of GaInAsP/InP DH lasers," IEEE J. Quantum. Electron., vol. 17, pp. 611- 619, 1981.

[74] C. S. Chang and S. L. Chuang, "Modeling of strained quantum-well lasers with spin-orbit coupling," IEEE J. Select. Topics Quantum Electron., vol. 1, p. 218, 1995.

[75] C. S. Chang, S. L. Chuang, J. R. Minch, W. W. Fang, Y. K. Chen, and T. Tanbun-Ek, "Amplified spontaneous emission spectroscopy in strained quantum-well lasers" IEEE J. Select. Topics Quantum Electron., vol. 1, p. 1100, 1995.

[76] H. Wenzel, G. Erbert, and P. M. Enders, " Improved theory of the refractive index change in quantum-well lasers," IEEE J. Select. Topics Quantum Electron., vol. 5, p. 637, 1999.

[77] J. C. Simon, "GaInAsP semiconductor laser amplifiers for single-mode fiber communications," IEEE/OSA J. Lightwave Tech., vol. 9, p. 1286, 1987.

[78] T. D. Visser, H. Blok, B. Demeulenaere, and D. Lenstra, "Confinement factors and gain in optical amplifiers," IEEE J. Quantum. Electron., vol. 33, p. 1763, 1997.

[79] H. J. S. Dorren, D. Lenstra, Y. Liu, M. T. Hill, and G. D. Khoe, "Nonlinear polarization rotation in semiconductor optical amplifiers: theory and application to all-optical flip-flop memories," IEEE J. Quantum. Electron., vol. 38, p. 141, 2003.

[80] R. J. Manning, A. Antonopoulos, R. L. Roux, and A. E. Kelly, "Experimental measurement of nonlinear polarisation rotation in semicondutor optical amplifiers," Electron. Lett., vol. 37, p. 229, 2001.

[81] X. Yang, D. Lenstra, G. D. Khoe, and H. J. S. Dorren, "Nonlinear polarization rotation induced by ultrashort optical pulses in a semiconductor optical amplifier," Opt. Commun., vol. 223, p. 169, 2003.

[82] A. L. Smirl, T. F. Boggess, B. S. Wherrett, G. P. Perryman, and A. Miller, "Picosecond optically induced anisotropic state filling in semiconductors," Phys. Rev. Lett., vol. 49, p. 933, 1982.

[83] J. L. Oudar, A. Migus, D. Hulin, G. Grillon, J. Etchepare, and A. Antonetti, "Femtosecond orientation relaxation of photoexcited carriers in GaAs," Phys. Rev. Lett., vol. 53, p. 384, 1984.

[84] M. T. Portella, J. Y. Bigot, R. W. Schoenlein, J. E. Cunningham, and C. V. Shank, "k-space carrier dynamics in GaAs," Appl. Phys. Lett., vol. 60, p. 2123, 1992.

[85] R. Binder, H. S. Kohler, M. B. Onitz, and N. Kwong, "Green's function description of momentum-orientation relaxation of photoexcited electron plasmas in semiconductors," Phys. Rev. B, vol. 55, p. 510, 1997.

[86] J. Mark and J. Mork, "Subpicosecond gain dynamcis in InGaAsP optical amplifier:Experiment and theory," Appl. Phys. Lett., vol. 61, p. 2281, 1992.

[87] J. Mork, J. Mark, and C. P. Seltzer, "Carrier heating in InGaAsP laser amplifier due to two-photon absorption," Appl. Phys. Lett., vol. 64, p. 2206, 1994.

[88] A. Mecozzi and J. Mork, "Theory of heterodyne pump-probe experiments with femtosecond pulses", J. Opt. Soc. Am. B, **13**, 2473, 1996.

[89] J. Mork and A. Mecozzi, "Response function for gain and refractive index dynamics in active semiconductor waveguides," Appl. Phys. Lett., vol. 65, p. 1736, 1994.

[90] J. Mork and A. Mecozzi, "Theory of the ultrafast optical response of active semiconductor waveguides," J. Opt. Soc. Am. B, vol. 13, p. 1803, 1996.

[91] Z. I. Fabian, J. D. Sarma, and S. D. Sarma, "Spintronics: Fundamentals and applications," Rev. Mod. Phys., vol. 76, p. 323, 2004.

[92] R. Fiederling, M. Kelm, G. Reuscher, W. Ossau, G. Schmidt, A. Waag, and L. W. Molenkamp, "Injection and detection of a spin-polarized current in a light-emitting diode " Nature, vol. 402, p. 787, 1999.

[93] H. J. Zhu, M. Ramsteiner, H. Kostial, M. Wassermeier, H. P. Schönherr, and K. H. Ploog, "Room-temperature spin injection from Fe to GaAs," Phys. Rev. Lett., vol. 87, p. 016601, 2001.

[94] G. Schmidt, "Concepts for spin injection into semiconductors-a review," J. Phys. D: Appl. Phys., vol. 38, p. R107, 2005.

[95] M. I. Dyakonov and V. I. Perel, Optical Orientation, F. Meier B. P. Zakharchenya ed. vol. . Amsterdam: Elsevier, 1984.

[96] S. Pfalz, R. Winkler, T. Nowitzki, D. Reuter, A. D. Wieck, D. Hagele, and M. Oestreich, "Optical orientation of electron spins in GaAs quantum wells," Phys. Rev. B, vol. 71, p. 165305, 2005.

[97] M. Agrawal and G. S. Solomon, "Quantum-well band structure effects on the emission polarization from a spin-polarized electron reservoir," Appl. Phys. Lett., vol. 85, p. 1820, 2004.

[98] B. T. Jonker, Y. D. Park, B. R. Bennett, H. D. Cheong, G. Kioseoglou, and A. Petrou, "Robust electrical spin injection into a semiconductor heterostructure," Phys. Rev. B, vol. 62, p. 8180, 2000.

[99] B. T. Jonker, A. T. Hanbicki, D. T. Pierce, and M. D. Stiles, "A comment on spin nomenclature for semiconductors and magnetic metals," J. Magn. Magn. Mater., vol. 277, p. 24, 2004.

[100] M. C. d. Oliveira and H. B. Sun, "Modeling of optical detection of spin-polarized carrier injection into light-emitting devices," Phys. Rev. B, vol. 69, p. 085322, 2004.

[101] J. W. Tomm, T. Q. Tien, and D. T. Cassidy, "Spectroscopic strain measurement methodology: Degree-of-polarization photoluminescence versus photocurrent spectroscopy," Appl. Phys. Lett., vol. 88, p. 133504, 2006.

[102] V. P. Labella, D. W. Bullock, Z. Ding, C. Emery, A. Venkatesan, W. F. Oliver, G. J. Salamo, P. M. Thibado, and M. Mortazavi, "Spatially resolved spin-injection probability for gallium arsenide " Science, vol. 292, p. 1518, 2001.

[103] D. T. Pierce and F. Meier, "Photoemission of spin-polarized electrons from GaAs," Phys. Rev. B, vol. 13, p. 5484, 1976.

[104] S. F. Alvarado and P. Renaud, "Observation of spin-polarized-electron tunneling from a ferromagnet into GaAs," Phys. Rev. Lett., vol. 68, p. 1387, 1992.

[105] J. L. Oudar, A. Migus, D. Hulin, G. Grillon, J. Etchepare, and A. Antonetti, "Femtosecond orientation relaxation of photoexcited carriers in GaAs," Phys. Rev. Lett., vol. 53, p. 384, 1984.

[106] M. T. Portella, J. Y. Bigot, R. W. Schoenlein, J. E. Cunningham, and C. V. Shank, "k-space carrier dynamics in GaAs," Appl. Phys. Lett., vol. 60, p. 2123, 1992.

[107] M. Wraback and H. Shen, "A femtosecond, polarization-sensitive optically addressed modulator based on virtual exciton effects in an anisotropically strained multiple quantum well," Appl. Phys. Lett., vol. 76, p. 1288, 2000.

[108] H. J. S. Dorren, X. Yang, D. Lenstra, H. de Waardt, G. D. Khoe, H. Ishikawa, H. Kawashima, and T. Hasama, "Ultrafast refractive-index dynamics in a multiquantum-well semiconductor optical amplifier," IEEE Photon. Tech. Lett., vol. 15, pp. 792-794, 2003.

[109] H. Ju, S. Zhang, D. Lenstra, H. de Waardt, E. Tangdiongga, G. Khoe, and H. Dorren, "SOA-based all-optical switch with subpicosecond full recovery," Opt. Exp., vol. 13, pp. 942-947, 2005.

[110] M. Yildirim, J. P. Prineas, E. J. Gansen, and A. L. Smirla, "A near-room-temperature all-optical polarization switch based on the excitation of spin-polarized "virtual" carriers in quantum wells," J. Appl. Phys., vol. 98, p. 063506, 2005.

[111] M. J. LaGasse, K. K. Anderson, C. A. Wang, H. A. Haus, and J. G. Fujimoto, "Femtosecond measurements of nonresonant nonlinear index in AlGaAs," Appl. Phys. Lett., vol. 56, pp. 417-419, 1990.

[112] C. T. Hultgren and E. P. Ippen, "Ultrafast refractive index dynamics in AlGaAs diode laser amplifiers," Appl. Phys. Lett., vol. 59, pp. 635-637, 1991.

[113] H. J. S. Dorren, A. K. Mishra, X. Yang, Z. Li, H. Ju, H. de Waardt, D. Khoe, and D. Lenstra, "All-optical switching and wavelength conversion based on ultrafast nonlinearities in semiconductor optical amplifiers," Japanese J. Appl. Phys., vol. 43, pp. 5731-5741, 2004.

[114] B. Zhao and A. Yariv, Quantum well semiconductor laser vol. I. London: Academic Press, 1999.

[115] J. Mork and A. Mecozzi, "Non-adiabatic effects in semiconductor wave guides," SPIE, vol. 3944, pp. 658-672, 2000.

[116] D. B. Stefan Schmitt-Rink, Volker Heuckeroth, Peter Thomas, Peter Haring, Gerd Maidorn, Huib Bakker, Karl Leo, Dai-Sik Kim and Jagdeep Shah, "Polarization dependence of heavy- and light-hole quantum beats," Phys. Rev. B, vol. 46, p. 10460, 1992.

[117] D. Bennhardt, P. Thomas, R. Eccleston, E. J. Mayer, and J. Kuhl, "Polarization dependence of four-wave-mixing signals in quantum wells," Phys. Rev. B, vol. 47, p. 13485, 1993.

[118] E. J. Mayer, G. O. Smith, V. Heuckeroth, J. Kuhl, K. Bott, A. Schulze, T. Meier, D. Bennhardt, S. W. Koch, P. Thomas, R. Hey, and K. Ploog, "Evidence of biexcitonic contributions to four-wave mixing in GaAs quantum wells," Phys. Rev. B, vol. 50, p. 14730, 1994.

[119] T. Aoki, G. Mohs, T. Ogasawara, R. Shimano, M. Kuwata-Gonokami, and A. Yamaguchi, "Polarization dependent quantum beats of homogeneously broadened excitons " Opt. Exp., vol. 1, pp. 364-369, 1997.

[120] T. Aoki, G. Mohs, M. Kuwata-Gonokami, and A. A. Yamaguchi, "Influence of Exciton-Exciton Interaction on Quantum Beats," Phys. Rev. Lett., vol. 82, p. 3108, 1999.

[121] D. J. Lovering, R. T. Phillips, G. J. Denton, and G. W. Smith, "Resonant generation of biexcitons in a GaAs quantum well," Phys. Rev. Lett., vol. 68, pp. 1880, 1992.

[122] H. P. Wagner, A. Schtz, W. Langbein, J. M. Hvam, and A. L. Smirl, "Interaction-induced effects in the nonlinear coherent response of quantum-well excitons," Phys. Rev. B, vol. 60, p. 4454, 1999.

[123] J. Ishi, H. Kunugita, K. Ema, T. Ban, and T. Kondo, "Influence of exciton-exciton interactions on frequency-mixing signals in a stable exciton-biexciton system," Phys. Rev. B, vol. 63, p. 73303 2001.

[124] K. Bott, O. Heller, D. Bennhardt, S. T. Cundiff, P. Thomas, E. J. Mayer, G. O. Smith, R. Eccleston, J. Kuhl, and K. Ploog, "Influence of exciton-exciton interactions on the coherent optical response in GaAs quantum wells," Phys. Rev. B, vol. 48, p. 17418, 1993.

[125] T. Saiki, M. Kuwata-Gonokami, T. Matsusue, and H. Sakaki, "Photon echo induced by two-exciton coherence in a GaAs quantum well," Phys. Rev. B, vol. 49, p. 7817, 1994.

[126] T. Saiki, M. K.-G. Ohkawa, and T. Mitsuyu, "Role of biexciton state in excitonic resonant nonlinearity in homoepitaxial ZnSe," Solid State Commun., vol. 95, pp. 679-683, 1995.

[127] W. Schäfer, D. S. Kim, J. Shah, T. C. Damen, J. E. Cunningham, K. W. Goossen, L. N. Pfeiffer, and K. Khler, "Femtosecond coherent fields induced by many-particle correlations in transient four-wave mixing," Phys. Rev. B, vol. 53, p. 16429, 1996.

[128] M. Lindberg and S. W. Koch, "Effective Bloch equations for semiconductors," Phys. Rev. B, vol. 38, p. 3342, 1992.

[129] J. Erland and I. Balslev, Phys. Rev. B 48 (3), R1765, 1993.

[130] W. Wang, K. Allaart, and D. Lenstra, "Correlation between electron spin and light circular polarization in strained semiconductors," Phys. Rev. B, vol. 74, p. 073201, 2006.

[131] M. O. Scully and M. S. Zubairy, Quantum Optics. London: Cambridge University Press, 1997.

Publications Included in Thesis

1. "Correlation between electron spin and light circular polarization in strained semiconductor" Wang W., Allaart K., and Lenstra D., *Phys. Rev. B* **74**, 073201, (2006).
   Selected by *Virtual J. of Nanoscale Sci.* & *Tech.*, **14**, 8, (2006).

2. "All-Optical Limiting Using Asymmetric Nonlinear Directional Couplers", Wang W., Wang Y., Allaart K., and Lenstra D., *Applied Physics B*, **83**, 623, (2006).

3. "Ultrashort rectangular optical pulse generation by nonlinear directional couplers", Wang W., Wang Y., Allaart K., and Lenstra D., *Optics Communication*, **253**, 164, (2005).

4. "Semiconductor optical amplifier gain anisotropy: confinement factor vs. material gain", Wang W., Allaart K., and Lenstra D., *Electronics Letters*, **40**, 1602, (2004).

5. "Semiconductor optical Bloch equation explain polarization dependent four wave mixing quantum beats in bulk semiconductor", Wang W., Zhang J., Allaart K., and Lenstra D. (Chapter 6, Preparing).

6. "The mechanisms of the ultrafast birefringence in semiconductors induced by photon excitations", Wang W., Allaart K., and Lenstra D. (Chapter 5, Preparing).

7. "Study of ultra-fast optical pulse propagation in a nonlinear directional coupler", Wang Y., and Wang W., *Applied Physics B*, **79**, 52, (2004).

8. ”A simple and efficient numerical method for nonlinear pulse propagation in N-core optical couplers”, Wang Y., and Wang W., *Photon. Tech. Lett.* **16**, 1077, (2004).

9. ”Coherence transfer for the polarization dependent four wave mixing quantum beats in the bulk semiconductor”, W. Wang, K. Allaart, and D. Lenstra, Frontiers in Optics 2006/Laser Science XXII, Rochester, New York, U.S.A., Oct. 8-12, (2006).

10. ”Investigate strain in semiconductor by the polarization correlation between the emission light and the electron spin” W. Wang, K. Allaart, and D. Lenstra, The 17th Quantum Electronics and Photonics Conference, Manchester, U.K., Sep. 4-7 (2006).

11. ”Ultrafast birefringence in semiconductor optical amplifier due to the carrier momentum orientation relaxation”, Wang W. ; Allaart K.; and Lenstra D., CLEO/EQEC, Munich, Germany, Jun. 12-17, (2005).

12. ”Ultrafast birefringence in semiconductor optical amplifier due to the dipole orientation relaxation”, Wang W., Allaart K., and Lenstra D., COST, Metz, France, (2005).

13. ”Gain anisotropy in a semiconductor optical amplifier confinement factors or material gain”, Wang W., Allaart K., and Lenstra D., the 6th International Conference on Transparent Optical Networks, Wroclaw, Poland, Jul. 4-8, (2004).

14. ”Modified semiconductor Bloch equations explain polarization dependent four wave mixing signal quantum beats in bulk semiconductors”, Wang W., Allaart K., and Lenstra D., 11th Annual Symposium of the IEEE/LEOS Benelux Chapter, Eindhoven, The Netherlands, Nov. 3-Dec. 1, (2006).

15. ”Pulse propagation in asymmetrical cascaded nonlinear directional couplers and limiting effects.” Wang W., Wang Y., Allaart K., and Lenstra D., 10th Annual Symposium of the IEEE/LEOS Benelux Chapter, Mons, Belgium, Dec. 1-2, (2005).

16. ”Polarization Characterized Band Model for Layered Semiconductors”, Wang W., Allaart K., and Lenstra D., 9th Annual Symposium of the IEEE /LEOS Benelux Chapter, Ghent, Belgium, Dec. 2-3, (2004).

Summary

This thesis deals with some specific problems concerning the processing of ultrashort optical pulses and their interaction with semiconductors, which are both academically interesting and technically important. It includes the investigation of the ultrashort optical pulse propagation and coupling dynamics in the nonlinear coupled waveguide, and the coherent and incoherent processes of the light-semiconductor interactions, therewith the arising nonlinearities, e.g. ultrafast birefringence and four wave mixing.

The subpicosecond optical pulse propagation in the coupled nonlinear waveguide is investigated in chapter 2. The new feature found is that the waveguide's optical switching characteristics do not depend on the pulse profile, like a soliton, but on the product of the dispersion length and the coupling coefficient. Understanding this feature, a new normalization is proposed for the coupled nonlinear Schrödinger equations, which are commonly used in the study of the coupling dynamics. This new normalization is proved to be highly efficient in the study of the ultrashort pulse coupling. To solve the coupled nonlinear Schrödinger equations, a numerical method is developed, in combination with the Split-Step Fourier method.

By investigating the pulse propagation in the cascaded nonlinear coupled waveguide, a method for the subpicosecond rectangular optical pulse generation is proposed. In a nonlinear coupled waveguide with one self-focusing and one self-defocusing channel, the limiting feature is demonstrated for both continuous wave and ultrashort pulses.

After the general theory of the semiconductor electrical and optical properties is described and discussed in chapter 3, we focus on the polarization dependence of optical transitions. It is shown that the polarization dependence of the semiconductor optical amplifier is much better reproduced if a weak valence band non-degeneracy is introduced in a heterostructure SOA than when it is simply attributed to the confinement factor difference.

In chapter 4, the correlation between the electron spin polarization and cir-

cular polarization of the emitted light is studied, which is important for the future ultrafast and nano-scale spintronics. We rigorously derive and prove the relation $P_{cir} = -\frac{1}{2}\cos\theta P_s$ in case of a bulk semiconductor, with θ the angle between the observation direction and the spin polarization direction. The correlation is independent of the electron wave vector $\vec{k}$. Also, we investigate the relation in the uniaxially strained bulk semiconductors. We notice that, in uniaxially strained bulk semiconductors, it is possible to achieve high spin polarization of the electrons in certain parts of the energy spectrum by circularly polarized light. We also suggest a method for detection of strain in semiconductors by measuring the circular polarization of its luminescence.

In chapter 5, the induced transient birefringence in semiconductors due to an ultrashort pulse is studied. We derive the momentum space anisotropic distribution of the electrons excited by photons and show that the corresponding birefringence is rather weak. Based on the Semiconductor Bloch Equations (SBE), the dipole polarization non-equilibrium in a photo-exited semiconductor, its influence on the response to the polarized electric field, and the resulting ultrashort birefringence is studied and shown to yield a high birefringence ratio.

In chapter 6, the polarization dependence of the coherent light-semiconductor interaction on an ultrashort time scale is studied within a model that emphasizes the coherence of the photo-excited electron-hole pair correlation. Based on this model, the polarization dependence of quantum beats in four wave mixing signal of semiconductors is studied.

Samenvatting

Dit proefschrift handelt over enkele specifieke problemen betreffende het produceren en manipuleren van ultrakorte optische pulsen en hun wisselwerking met halfgeleiders. Problemen die zowel academisch interessant zijn als ook van technisch belang. Het proefschrift omvat een studie van de propagatie en koppelingsdynamica in niet-lineaire gekoppelde golfgeleiders en eveneens van de coherente en incoherente processen bij interactie van licht met halfgeleiders en de daarbij optredende niet-lineaire verschijnselen, zoals ultrasnelle dubbele breking en "four-wave mixing" (FWM), waarbij twee laserpulsen in de halfgeleider een derde lichtbundel genereren.

De propagatie van optische pulsen korter dan een picoseconde in gekoppelde niet-lineaire golfgeleiders werd onderzocht in hoofdstuk 2. Als nieuw aspect hiervan werd gevonden dat het karateristieke gedrag als optische schakelaar niet afhangt van het profiel van de pulsen, zoals bij een soliton, maar van het product van de dispersielengte en de koppelingscoëfficient. Dit verschijnsel begrepen hebbende, werd een nieuwe normering voorgesteld voor de niet-lineaire Schrödinger vergelijkingen, die gewoonlijk worden toegepast bij het bestuderen van de koppelingsdynamica. Deze nieuwe nornering blijkt zeer efficiënt bij het bestuderen van de koppeling tussen ultrakorte pulsen. Om de gekoppelde niet-lineaire Schrödinger vergelijkingen op te lossen werd een numerieke methode ontwikkeld, in combinatie met de Split-Step Fourier methode.

Door de propagatie van pulsen in een cascade van niet-lineaire gekoppelde golfgeleiders te onderzoeken, werd een methode voor het genereren van rechthoekige optische pulsen voorgesteld. Eveneens werd aangetoond dat een systeem van gekoppelde golfgeleiders, waarvan de ene zelf-focusserend en de andere zelf-defocusserend is, kan dienen als begrenzer van de pulssterkte, zowel bij continue golven als bij ultrakorte pulsen.

Na de beschrijving van algemene electronische en optische eigenschappen van de halfgeleider in hoofdstuk 3, wordt speciaal aandacht besteed aan de polarisatieafhankelijkheid van de optische overgangen. Getoond wordt dat de polarisatieafhankelijkeid van de halfgeleider optische versterker (SOA) veel beter wordt beschreven wanneer men een zwakke niet-ontaarding van de valentieband in een SOA met gelaagde samenstelling ("heterostructure SOA") aanneemt dan wanneer men die afhankelijkheid toeschrijft aan het verschil in opsluiting ("confinement") voor verschillende polarisaties.

In hoofdstuk 4 wordt de correlatie tussen de polarisatie van de electronspin en de circulaire polarisatie van het uitgezonden licht bestudeerd. Deze is van belang voor de ultrasnelle en nano "spintronica". Een stricte afleiding en bewijs wordt gegeven voor de relatie $P_{cir} = -\frac{1}{2}P_s\cos\theta$ voor het geval van een bulk halfgeleider, waarin θ de hoek is tussen de richting van de waarnemer en die van de spinpolarisatie. De correlatie hangt niet af van de golfvector $\vec{k}$ van het electron. De relatie tussen de polarisatie van de spin en de circulaire polarisatie van het licht is ook onderzocht voor bulk halfgeleiders waarin een axiaal symmetrisch spanningsveld bestaat. We zien dan dat het in dat geval mogelijk is om met circulair gepolariseerd licht, in bepaalde delen van het energiespectrum, een hoge graad van spinpolarisatie te bereiken. Ook stellen we voor de eventueeel aanwezige spanning te detecteren door meting van de circulaire polarisatie van de luminiscentie.

In hoofdstuk 5 wordt de kortstondige, door een ultrakorte puls veroorzaakte, dubbele breking in halfgeleiders bestudeerd. We geven een afleiding van de anisotrope verdeling in de impulsruimte van de door de fotonen geëxciteerde electronen en tonen aan dat de daardoor veroorzaakte mate van dubbele breking vrij zwak is. Gebaseerd op de Bloch vergelijkingen voor de halfgeleider (Semiconductor Bloch Equations; SBE) worden de niet-evenwichts dipoolpolarisatie in een foto-geëxciteerde halfgeleider, de invloed daarvan op de response na een gepolariseerd electrisch veld en de daaruit voortvloeiende kortdurende dubbele breking bestudeerd. Getoond wordt dat een hoge mate van dubbele breking voor kan komen.

In hoofstuk 6 wordt de polarisatieafhankelijkheid van de coherente wisselwerking tussen licht en een halfgeleider op een ultrakorte tijdschaal bestudeerd, met een model dat de coherente correlatie van de geëxciteerde electron-gat paren benadrukt. In dit opzicht verschilt het van de gebruikelijke SBE. Gebaseerd op dit model wordt de polarisatieafhankelijkheid van de zwevingen ("quantum beats") in het gedetecteerde FWM signaal uit de halfgeleider afgeleid.

Acknowledgements

The thesis is the completion of my three years PhD study at the Vrije Universiteit Amsterdam. I would like to take this opportunity to thank all of the people who contributed directly or indirectly to this thesis.

First there are my supervisors Klaas Allaart and Daan Lenstra. I particularly appreciate Daan for his encouraging my fantastic ideas, the full freedom to follow my research interests, and the great guidance. Also I learned lots from his great personality, which would be a wealth in my future life. Klaas helped me a lot during my transfer from engineer to physics scope, I particularly appreciate his great patience in teaching me, hundreds of instructive discussions, and fruitful work. Without these, I would have hardly finished the study. In a word, the combination of both supervisors substantially contributed to my successful research work.

I also would like to thank the members of the reading committee: Alfred Driessen, Andrei Silov, Wim Ubachs and Taco Visser for taking time to read my thesis carefully and give some useful comments.

Furthermore I would like to thank Zhonggui Li, Shaoxian Zhang, Yong Liu, Jianming Zeng in Eindhoven Univ. of Tech, Wei Yan in Salford University, and the members of the Theoretical Physics Group at Vrije Universiteit Amsterdam. Discussing with you was always inspiring. Particular I would like to thank Youfa Wang in Avago, Singapore, for his guiding during my hardest time in these three years.

Moreover I would like to thank the Chinese community in Amsterdam. With these friends, I have a wonderful life in this fantastic city.

Finally I would like to thank my wife Jun for her support, encouragement and patience, also my parents and other relatives in China.

Curriculum Vitae

Wenfeng Wang was born in Tianmen, Hubei Province, China, in 1978. He received his B. Eng. degree in electrical material physics in 2000 from the Physics Department, Hubei University, China. From September 2000 to July 2001, he worked as an engineer in a high-tech company in China. In 2001, he started to study in COM research center, Denmark Technology University(DTU), with the financial support from Photonics Crystal Fiber company and Danish government. From September 2002 to July 2003, he conducted his master thesis at OFS (former Lucent, Denmark) under a cooperative project with COM. For this project he worked on the Raman fiber amplifier based on the dispersion compensation fiber. He received his M. Eng. degree in optical telecommunications in 2003 from DTU.

In October 2003 he started working toward a Ph.D. degree in the Quantum Electronics Group of the Physics Department at Vrije Universiteit Amsterdam, the Netherlands. The study is supported by the Freeband "Towards Ultrafast Communications (TUC)" project of the Netherlands Ministry of Economic Affairs.